

普通高等教育“十一五”国家级规划教材

全国电子信息类优秀教材一等奖

“十二五”江苏省高等学校重点教材

江苏省精品教材

电子信息科学与工程类专业规划教材

数字图像处理

(第3版)

胡学龙 编著



电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书是普通高等教育“十一五”国家级规划教材、全国电子信息类优秀教材一等奖教材、“十二五”江苏省高等学校重点教材（编号：2013-1-052）、江苏省精品教材、江苏省研究生优秀课程的参考书。本书强调现代数字图像处理理论与应用的紧密结合，在阐述基本原理的基础上，通过习题、实验和计算机软件工具介绍实现图像处理的基本方法。全书共 11 章，主要内容包括数字图像的发展简史、基本图像处理系统、图像数字化与显示、图像变换与二维数字滤波、图像编码与压缩、图像增强、图像复原、图像分割、形态学图像处理、彩色图像处理、图像的数字水印、图像分析系统、基于内容的检索、图像存档与通信系统、图像融合、图像处理软件的开发等。本书配套电子课件、习题参考答案、课程网站和以 MOOC 的形态和标准制作的微视频。

本书可作为高等学校电子信息工程、电子信息科学与技术、通信工程、计算机科学与技术、自动化、生物医学工程等专业高年级本科生教材，也可作为相关专业本科生和研究生的教学参考书，还可供相关工程技术人员学习参考。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

数字图像处理/胡学龙编著. —3 版. —北京：电子工业出版社，2014.8

电子信息科学与工程类专业规划教材

ISBN 978-7-121-23887-1

I. ①数… II. ①胡… III. ①数字图像处理—高等学校—教材 IV. ①TN911.73

中国版本图书馆 CIP 数据核字（2014）第 169296 号

策划编辑：王羽佳

责任编辑：王羽佳 特约编辑：曹剑锋

印 刷：

装 订：

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×1092 1/16 印张：16.5 字数：477 千字

版 次：2014 年 8 月第 1 版

印 次：2014 年 8 月第 1 次印刷

印 数：4000 册 定价：35.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前 言

在人类获取外界信息的渠道中,由眼睛感受的视觉信息占据举足轻重的地位。除了可视的自然景象外,宏观和微观影像及大量由计算机生成的图像丰富了视觉的范畴,成为多媒体世界最重要的成员之一。图像处理技术产生于 20 世纪初,其发展速度之快、应用范围之广、与其他学科交叉之紧密,令人瞩目。它在科学技术、国民经济的各个领域以及人们的工作和生活中发挥越来越重要的作用。

在 21 世纪,数字图像处理正以强劲的发展态势与后劲,朝着智能化、网络化、个人化、实时化等方向发展,图像信息已成为第四代移动通信、物联网、大数据和云计算技术研究的重要对象。近年来,平板电脑和大屏智能手机普及,其高清拍照、在线视频、二维码识别、名片识别、视频聊天等功能深受大众欢迎。同时图像处理技术在我国的基础研究与工程技术的振兴中做出了显著贡献,特别在遥感遥测和宇航工程方面更是立下了汗马功劳!“高分一号”卫星作为我国高分辨率对地观测系统首发星的成功发射开启了我国对地观测的新时代。以近年来嫦娥工程为例,“嫦娥一号”和“嫦娥二号”卫星所携带的干涉成像光谱仪、激光高度计和 CCD 立体相机共同完成了月球表面三维立体影像的获取工作。配备了降落相机的“嫦娥三号”在月面上实现了软着陆,着陆器安装了地形地貌相机、近紫外月基天文望远镜、极紫外相机等探测设备。月球车配备全景相机等成像设备,在月球表面进行巡视探测。在微观世界,我国科学家在水科学领域取得重大突破,首次拍摄到了水分子的内部结构,使得在实验中直接解析水的氢键网络构型成为可能。

在本科阶段“数字图像处理”作为电子信息工程、电子信息科学与技术、通信工程、计算机科学与技术、自动化等专业的专业课,对提高学生认识世界和改造世界的能力大有裨益。虽然各专业对图像的获取、变换、编码与压缩、增强与复原、分割与目标识别等处理及硬件与软件的实现要求不尽相同,学生的兴趣点也有差异,但基本理论、基础知识、基本技能对于不同专业都是不可缺少的。

图像处理具有鲜明的学科交叉性。学习图像处理涉及众多知识领域。获取图像,要了解成像原理和光传感器的结构,这涉及光学和电子学基础;为了表示图像,要用到数学分析、矩阵论、随机过程;对图像进行增强,要了解视觉的生理和心理特点;对图像进行压缩,必须了解信息论和编码的基本理论和方法;在计算机中实际进行图像处理时,对编程语言和软件工具 MATLAB、C++ 和 Photoshop 等也应该熟悉。现代图像处理技术更是将最新的理论和方法与本学科紧密结合,如数学形态学、人工神经网络(ANN)、模糊集与粗糙集理论、遗传算法等演化算法、人工生命等复杂性系统等,这些技术的引入使图像处理更加智能化。

本课程作为专业课,学时较少,但承接前面课程的概念较多,新概念也较多。近年来,本科毕业设计和研究生在图像处理、图像分析和机器视觉等方面的研究课题日趋增多。该课程产生的延伸影响越来越大。如何有效地组织教材,使之成为一个有机的整体,是摆在我们面前的难题。20 多年来,对图像处理技术的学习,使我们感到图像处理技术博大精深。我们将教材编写视为创新性的工作,尽管国内外类似的教材较多,但我们希望在立足于吸收众多专家成果的基础上,反映我们对图像处理学习、讲授、指导研究生和从事科研实践的体会和感受。严谨、求实、创新,是我们编写本书的基本原则。

在推进研究性教学、卓越工程师培养计划和 MOOC(大型开放式网络课程)教学的时代背景下,对第 2 版教材进行了全面修订,进一步提高教材的可读性、实际教学效果,并为研究性图像处理教学奠定坚实基础,注重图像处理理论和技术的实际应用,做到与时俱进,精益求精。第 3 版特别对原理

的验证、思路的拓展进行了深入的研究。新教材进一步完善了本学科的知识体系,按照学生的学习规律,在循序渐进的指导原则下,对部分章节重新编排,增加了不少可供实际操作的程序,对实验内容、要求和步骤做了较大的更新,对习题进行了必要的增补和修订。为了便于教学,我们在书后列出了在双语教学实践中选取的图像处理专业英文词汇和 MATLAB 图像处理工具箱。本教材按照 48 学时(含实验)编排,各校、各专业可根据人才培养方案,对学时和内容进行编排。

本书的第 1、2 版在读者和专家们的关爱下,获得了良好的社会声誉:第 1 版 2006 年被评为普通高等教育“十一五”国家级规划教材,2007 年被评为江苏省精品教材,2010 年又成为江苏省研究生优秀课程“图像信号处理”的主要参考书,2012 年荣获全国电子信息类优秀教材(研究生、本科生教育优秀教材)一等奖。由中国电子教育学会、电子工业出版社和扬州大学共同举办的“数字图像处理”及信号处理类课程教师培训与研修班将本教材作为培训教材,作者围绕图像处理学科前沿和课程改革、研究性教学、实践教学等内容做了专题报告。同时,上百所高等学校的使用也在不断推动着第 3 版的问世。2013 年本书被评为江苏省“十二五”重点教材立项再版建设项目。

本教材的多媒体互动教学资源库建设入选国家新闻出版改革发展项目库 2013 年度项目。为丰富教材内容和采用本教材的师生提供增值服务,本书提供免费的数字资源,包括电子课件、习题参考答案和配套课程网站,并以 MOOC 的形态和标准配套了微视频。请登录华信教育资源网(<http://www.hxedu.com.cn>)注册下载。

本书的修订工作由扬州大学胡学龙教授完成,孙磊、陈舒涵等老师,陆慧敏、范谦、周金龙、颜子越、陆李娜等几位博硕士生也参与了修订工作。本书的修订工作得到了扬州大学、江苏省高等学校现代教育技术研究会各级领导的支持。在与日本九州工业大学芹川圣一(Seiichi SERIKAWA)教授联合培养研究生的过程中,本书的内容得到了不断充实。电子工业出版社各级领导和相关编辑的精心组织、细心审阅和修改保证了本书高质量地如期出版。书中还参考了大量国内外期刊、专著、教材和图片。在此一并表示衷心的感谢!

本书的修订工作由扬州大学胡学龙教授完成,孙磊、陈舒涵等老师,陆慧敏、范谦、周金龙、颜子越、陆李娜等几位博士生硕士生也参与了修订工作。本书的修订工作得到了扬州大学、江苏省高等学校现代教育技术研究会各级领导的支持。在与日本九州工业大学芹川圣一(Seiichi SERIKAWA)教授联合培养研究生的过程中,本书的内容得到了不断充实。电子工业出版社各级领导和相关编辑的精心组织、细心审阅和修改保证了本书高质量地如期出版。书中还参考了大量国内外期刊、专著、教材和图片。在此一并表示衷心的感谢!

本书的编著工作得到扬州大学出版基金、江苏省第七批“六大人才高峰”高层次人才项目、国家自然科学基金资助项目(61301220、51108402)、江苏省卓越工程师培养计划(软件类)试点项目、江苏省高等教育教改研究重中之重课题(2011JSJG012)及重点课题(2013JSJG069)、扬州大学教改重点课题(YZUJX2014-9A、YZUJX2014-56D)等项目的资助。

由于作者的教学和科研水平有限,书中难免有不当之处,敬请同行专家和读者不吝指正。同行专家、读者的宝贵意见将给作者更大的激励,并有助于提高以后面世的新版本的内容质量和进一步提高该课程的教学质量。读者意见请反馈至 wylj@phei.com.cn。

胡学龙
2014 年 8 月

目 录

第 1 章 绪论	1	3.3 图像输入/输出设备	42
1.1 图像、像素及数字图像处理	2	3.3.1 图像输入设备	42
1.2 数字图像处理发展简史	2	3.3.2 图像输出设备	47
1.3 图像处理的目的、任务与特点	6	3.4 实验：图像的数字化	49
1.3.1 图像处理的目的	6	本章小结	50
1.3.2 图像处理的任务	7	思考题与习题	50
1.3.3 数字图像处理的特点	11	第 4 章 图像变换与二维数字滤波	51
1.4 数字图像处理的应用	12	4.1 二维离散傅里叶变换 (DFT)	52
本章小结	15	4.1.1 二维连续傅里叶变换	52
思考题与习题	15	4.1.2 二维离散傅里叶变换	53
第 2 章 图像处理基本知识	16	4.1.3 二维离散傅里叶变换的性质	54
2.1 人类视觉与色度学基础	17	4.2 二维离散余弦变换 (DCT)	57
2.1.1 人类的基本视觉特性	17	4.2.1 一维离散余弦变换	57
2.1.2 三基色原理	17	4.2.2 二维离散余弦变换	58
2.1.3 光度学基本知识	18	4.2.3 二维 DCT 的应用	58
2.2 连续图像的数学描述	19	4.3 二维离散沃尔什-哈达玛变换	
2.3 基本的图像处理系统	20	(DHT)	59
2.3.1 图像处理硬件	20	4.3.1 沃尔什变换	59
2.3.2 图像处理软件	23	4.3.2 哈达玛变换	61
2.4 图像的统计特征	28	4.4 卡胡南-列夫变换 (K-L 变换)	63
2.5 实验：图像的基本操作和基本统计		4.5 二维离散小波变换	64
指标计算	30	4.5.1 小波分析的思想来源	65
本章小结	32	4.5.2 连续小波变换	66
思考题与习题	33	4.5.3 一维离散小波变换	67
第 3 章 图像的数字化与显示	34	4.5.4 二维离散小波变换	67
3.1 图像数字化的基本过程	35	4.6 二维数字滤波器	71
3.1.1 图像的采样和量化	35	4.7 实验：图像变换与二维数字滤波	72
3.1.2 数字图像的数据量	37	本章小结	74
3.1.3 采样和量化参数的选择	37	思考题与习题	74
3.1.4 二维采样定理	39	第 5 章 图像编码与压缩	76
3.2 图像的量化方法	40	5.1 概述	77
3.2.1 量化方法分类	40	5.1.1 数据压缩的基本概念	77
3.2.2 标量化	40	5.1.2 图像编码压缩的必要性	78
3.2.3 向量量化	41	5.1.3 图像编码压缩的可能性	78

5.1.4	图像编码压缩的技术指标	79	6.4.3	中值滤波	118
5.1.5	数据压缩方法的分类	81	6.4.4	边界保持类滤波	119
5.2	统计编码	81	6.5	锐化	120
5.2.1	Huffman 编码	81	6.5.1	空间域差分法	120
5.2.2	算术编码	83	6.5.2	频率域高通滤波法	125
5.3	预测编码	85	6.6	实验: 图像增强	125
5.3.1	预测编码基本原理	85	本章小结		127
5.3.2	线性预测编码	86	思考题与习题		128
5.3.3	自适应预测编码	87	第 7 章 图像复原		130
5.4	变换编码	87	7.1	图像退化原因与复原技术分类	131
5.4.1	变换编码的基本原理	87	7.1.1	连续图像退化的数学模型	131
5.4.2	变换编码的系统结构	88	7.1.2	离散图像退化的数学模型	133
5.4.3	变换编码方案的选取	88	7.2	逆滤波复原	134
5.4.4	整数小波变换与图像压缩	90	7.3	约束复原	135
5.5	二值图像编码	92	7.3.1	约束复原的基本原理	135
5.5.1	跳跃空白编码	92	7.3.2	维纳滤波方法	135
5.5.2	游程长度编码	92	7.3.3	平滑度约束最小平方滤波	137
5.6	图像压缩编码标准	93	7.4	非线性复原方法	139
5.6.1	彩色与灰度图像压缩标准 JPEG	93	7.4.1	最大后验复原	140
5.6.2	二值图像压缩标准 JBIG	95	7.4.2	最大熵复原	140
5.6.3	JPEG2000 静态图像压缩标准	97	7.4.3	投影复原	141
5.7	实验: 图像编码与压缩	98	7.4.4	同态滤波复原	142
本章小结		100	7.5	盲图像复原	143
思考题与习题		101	7.5.1	直接测量法	143
第 6 章 图像增强		102	7.5.2	间接估计法	143
6.1	概述	103	7.6	几何失真校正	146
6.1.1	图像增强的目的	103	7.6.1	典型的几何失真	146
6.1.2	图像增强技术的分类	103	7.6.2	空间几何坐标变换	146
6.1.3	直方图的概念	104	7.6.3	校正空间像素点灰度值的确定	147
6.2	灰度修正	105	7.6.4	鱼眼图像校正方法简介	150
6.2.1	灰度级校正	106	7.7	图像修复技术简介	151
6.2.2	灰度变换	107	7.8	实验: 图像复原	152
6.2.3	灰度直方图变换	110	本章小结		153
6.3	同态增晰	113	思考题与习题		154
6.3.1	问题的由来	113	第 8 章 图像分割		155
6.3.2	增晰原理	113	8.1	概述	156
6.3.3	增晰算法	113	8.1.1	图像分割的目的和任务	156
6.4	平滑	114	8.1.2	图像分割的集合定义	156
6.4.1	图像噪声	114	8.1.3	图像分割的分类	157
6.4.2	邻域平均法	115	8.2	像素的邻域和连通性	157

8.3	图像的阈值分割技术	159	9.4	实验: 形态学图像处理	188
8.3.1	基本原理	159	本章小结	190	
8.3.2	全局阈值分割	160	思考题与习题	190	
8.3.3	局部阈值分割	163	第 10 章 彩色图像处理	191	
8.4	图像的边缘检测	163	10.1	彩色图像处理的基本问题	192
8.4.1	边缘检测的基本原理	163	10.2	颜色空间的表示及其转换	192
8.4.2	梯度算子	164	10.2.1	RGB 模型	192
8.4.3	拉普拉斯算子	165	10.2.2	Munsell 模型	193
8.4.4	拉普拉斯-高斯算子	166	10.2.3	HSV 模型	193
8.4.5	坎尼边缘检测算子	166	10.2.4	HSI 模型	194
8.4.6	方向算子	167	10.2.5	YUV 模型	194
8.4.7	边缘跟踪	167	10.2.6	CMYK 模型	195
8.5	霍夫变换	169	10.2.7	$L^*a^*b^*$ 模型	195
8.5.1	直角坐标系中的霍夫变换	170	10.2.8	RGB 与 HSV 空间的相互 转换	196
8.5.2	极坐标系中的霍夫变换	171	10.2.9	RGB 与 YUV 空间的相互 转换	197
8.6	区域生长法	171	10.2.10	RGB 与 HSI 空间的相互 转换	198
8.7	图像分割方法的比较	173	10.2.11	真彩色、索引色和灰度 图像转换成二值图像	199
8.7.1	边缘检测的优缺点	173	10.3	颜色空间的量化	200
8.7.2	区域分割的优缺点	174	10.4	抖动技术	201
8.8	实验: 图像分割	174	10.5	假彩色处理	201
本章小结	175	10.6	彩色图像增强	202	
思考题与习题	175	10.6.1	真彩色增强	202	
第 9 章 形态学图像处理	177	10.6.2	伪彩色增强	203	
9.1	概述	178	10.7	彩色图像形态学	205
9.1.1	数学形态学的发展简史及基本 思想	178	10.7.1	彩色图像形态学的基本方法	205
9.1.2	几个基本概念	179	10.7.2	基于数学形态学的彩色图像 滤波	206
9.2	二值形态学	180	10.8	实验: 彩色图像处理	207
9.2.1	二值腐蚀	181	本章小结	209	
9.2.2	二值膨胀	181	思考题与习题	210	
9.2.3	二值开运算	183	第 11 章 数字图像处理的应用	211	
9.2.4	二值闭运算	184	11.1	图像处理在数字水印上的应用	212
9.3	灰值形态学	184	11.1.1	概述	212
9.3.1	灰值腐蚀	184	11.1.2	数字水印的衡量标准	212
9.3.2	灰值膨胀	186	11.1.3	数字水印的分类	212
9.3.3	灰值开运算	186	11.1.4	实现数字水印的一般步骤	213
9.3.4	灰值闭运算	186			
9.3.5	灰值形态学梯度	187			
9.3.6	高帽变换和低帽变换	188			
9.3.7	开-闭运算和闭-开运算	188			

11.1.5	图像水印举例	213	11.4.3	主要解决的问题和技术要点	229
11.2	基于数学形态学的图像颗粒度分析系统	215	11.4.4	DICOM 图像格式	231
11.2.1	概述	215	11.4.5	DICOM 3.0 标准及其面向对象的实现	231
11.2.2	求图像中目标的面积和颗粒度	215	11.4.6	小结	234
11.2.3	实验结果与分析	217	11.5	基于多分辨率分析的图像融合方法	234
11.2.4	小结	218	11.5.1	图像融合的层次	234
11.3	基于内容的图像检索 (CBIR)	218	11.5.2	图像融合的方法	235
11.3.1	概述	218	11.6	数字图像处理的发展趋势	237
11.3.2	基于内容图像检索的发展	219	11.7	基于 MATLAB GUI 图像处理的软件的开发	239
11.3.3	图像特征的概念	219	11.7.1	GUIDE 基本操作	239
11.3.4	基于内容的图像检索系统的框架	220	11.7.2	GUIDE 图像处理软件设计实例	239
11.3.5	相似度测量公式	222	11.8	实验: 基于 GUIDE 的图像处理软件开发	243
11.3.6	基于内容的图像检索系统简介	223	本章小结	244	
11.3.7	基于内容的图像检索技术的研究热点	224	思考题与习题	244	
11.3.8	一种基于颜色和纹理特征的彩色图像检索算法	224	附录 A	常用词汇中英文对照表	245
11.4	数字化医院中的图像存档与通信系统 (PACS)	227	附录 B	常用 MATLAB 图像处理工具箱函数	251
11.4.1	概述	227	参考文献	256	
11.4.2	国内外发展现状	228			

第1章 绪 论

➡ 内容提要

本章介绍数字图像的基本概念、数字图像处理的发展简史、数字图像处理的任务与特点、数字图像处理的应用。

➡ 知识要点

- 图像、数字图像、像素。
- 图像处理、数字图像处理。
- 数字图像处理的目的、任务与特点。
- 数字图像处理的典型应用。

➡ 教学建议

- 本章教学安排2学时，对全书的学习有一定的导读作用。
- 学习本课程的先修知识主要包括：高等数学、物理学（主要包括电磁学、光学等）、线性代数、概率论与随机过程、数字信号处理、信息论与编码、微机原理、软件技术基础等。
- 从知识体系看，本课程的后续课程包括图像通信、多媒体技术等。在研究生阶段，可进一步扩展至图像分析、视频信号处理、模式识别等课程。
- 重点了解图像处理的任务，明确本课程与所学专业的联系，学习和研究数字图像处理的基本方法。

1.1 图像、像素及数字图像处理

随着信息化进程的不断推进，我们已经进入大数据（big data）和云计算（cloud computing）时代。“大数据”是由数量巨大、结构复杂、类型众多数据构成的数据集合，是基于云计算的数据处理与应用模式。人们普遍认为大数据具有4V特征：①数据量（Volume）巨大，从TB级别跃升到PB级别；②时效性（Velocity）要求高速，处理速度快，常常需秒级完成；③多样性（Variety）明显，体现在数据类型繁多，包括互联网文本和文件、网络日志、社会数据、电子商务、视频、图片等；④精确性（Veracity）反映数据的价值密度较低，当数据巨量、来源多元时，这些数据本身的质量存在可疑型，当然数据分析后的利用价值可想而知。如在监控过程中形成的TB级数据中，能够作为取证的有用数据可能发生在秒数量级。由于图像和视频数据已经占到整个大数据的80%以上。这些数据如何表示、采集、处理、传输、管理和利用是我们迫切要解决的问题。人们希望通过图像等数据的整合共享，交叉复用，形成智力资源和知识服务能力。

图像（image）是自然界景物的客观反映。自然界的图像无论是在亮度、色彩，还是在空间分布上，其数值都是以模拟形式出现的。图1.1是用普通胶片照相机拍摄的一幅自然景物图像。传统电视等系统摄取、传输和显示的图像是模拟图像，借助于连续信号处理理论分析、设计、测试和存储图像，无法采用数字计算机或其他数字信号处理系统进行处理、传输和存储。

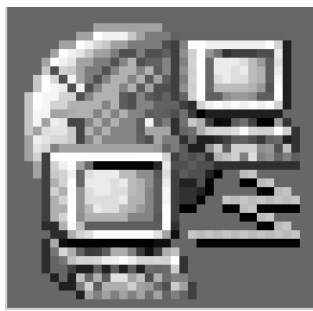
在数字图像领域，我们将图像视为由许多大小相同、形状一致的像素（picture element，简称pixel）组成。因此，一幅图像可以用二维矩阵加以表示。图1.2所示为一个小图标及该图标放大4倍后的方形像素。这样，数字图像可以用矩阵表示。图像的数字化包括采样（sampling）、量化（quantization）和编码（coding）三个主要步骤。在空间对连续坐标进行离散化的过程称为采样，而进一步将图像的幅度值（可能是灰度或色彩）整数化的过程称为量化。编码则是按照一定的规律，把量化后的值用二进制数字表示。这样得到的数字图像信号可以通过电缆、微波干线、卫星通道等数字线路传输。在接收端则与上述数字化过程相反恢复成原来的模拟图像信号。上述数字化的过程又称为脉冲编码调制（PCM）。一般情况下，一个完整的图像处理系统输入和显示的都是便于人眼观察的模拟图像。



图 1.1 自然景物图像



(a) 原图



(b) 将原图放大4倍

图 1.2 像素

数字图像处理（digital image processing）是用计算机或数字设备对图像进行加工、分析，以达到所需效果的技术。目前，图像处理一般指数字图像处理。

1.2 数字图像处理发展简史

远古时代人们对外界的感受是直观的，象形文字就是用视觉印象表达抽象意义的一种表达形式。

望远镜延伸了人的宏观视觉范围,而显微镜帮助人们洞察微观世界。19世纪中叶,照相机的发明使人们对观察的印象成为永恒的记录;而上世纪,彩色胶卷的问世实现了从无色的灰色记录到真实的彩色记录。

20世纪20年代,图像处理首次采用图像压缩技术应用于改善伦敦和纽约之间海底电缆发送的图片质量。离散数学的创立和完善,为数字图像处理奠定了理论基础。1946年数字计算机的出现使图像的获取、处理、传输和存储产生了质的飞跃,也使“数字图像处理”几乎成了“计算机图像处理”的代名词,成为当代图像处理的主流。数字图像处理最早出现于20世纪50年代,当时的电子计算机已经发展到一定水平,人们开始利用计算机来处理图形和图像信息。早期的计算机在计算速度、存储容量和软件处理功能等主要方面,难以满足对图像数据进行实时处理的要求。随着计算机软硬件技术的迅速发展,计算机处理图像的性能有了大幅度的提高。过去只能用大型计算机完成的处理功能,现在个人计算机(PC)上就能够方便地实现。

数字图像处理作为一门学科,大约形成于20世纪60年代初期。早期的图像处理的目的是提高图像的质量,以改善人的视觉效果为目的。常用的图像处理方法有图像增强、复原、编码、压缩等。数字图像处理首次成功地应用在1964年美国宇航局喷气推进实验室(NASA JPL),当时对“徘徊者7号”探测器发来的几千张月球照片进行了几何校正、灰度变换、去除噪声等处理,并考虑了太阳位置和月球环境的影响,用计算机绘制了月球表面的照片。随后又对探测飞船发回的近十万张照片进行更为复杂的图像处理,获得了月球的地形图、彩色图及全景镶嵌图,为人类登月创举奠定了坚实的基础。在以后的宇航空间技术,如对火星、土星等星球的探测研究中,数字图像处理技术都发挥了巨大的作用。直到现在,图像处理在航天技术领域还是不可缺少的重要手段。1990年发射的轨道高度约600 km的哈勃太空望远镜能辨别130亿光年以外的物体,人们甚至通过它观察到由于大爆炸形成时宇宙初期的奇形怪状的星体。2009年NASA“火星科学实验室(MSL)”探索计划(又名“好奇号”火星车)被视为与哈勃项目相当的“极为重要的旗舰项目”。2012年8月新型火星探测器“好奇号”着陆火星表面。“好奇号”装载了一系列成像设备(见图1.3)。① 主相机(MastCam):桅杆上的主要成像工具,一个中焦段定焦、一个望远定焦,负责拍摄火星地貌的1600×1200 CCD高解析度彩色照片和720P每秒10帧视频。② 火星手持透镜成像仪(MAHLI):安装在“好奇号”的机械臂末端,可以拍摄小到只有12.5微米的地貌特征彩色照片。③ 火星降落成像仪(MARDI):是一台安装在主车腹部的小型摄影机,离火星地表附近时启动,负责拍摄火星车降落地面过程中每秒5帧的影像。④ 化学相机(ChemCam):可以向约9米外的火星岩石发射高能激光,而后分析蒸发的岩石成分。⑤ 导航相机(NavCam):装在桅杆上每个有45度的视野的两对导航用的黑白3D相机,图1.3中未画出。⑥ 避险相机(HazCam):在四个角落的较低位置各装有每个约有120度视野的一对避开障碍用的黑白3D相机,图1.3中未画出。

图像处理技术从航天技术扩展到遥感遥测、生物医学、军事和公安、工业机器人、人民生活等诸多方面。美国陆地卫星(Landsat)5号TM(Thematic Mapper,主题测绘仪、主题成像传感器,一种多波段扫描型地球资源敏感仪器)图像,可从可见光(蓝绿、绿红、红)分布到红外(近红外、热红外)共7个光谱波段。所采用的全球参考系统以轨道号(PATH/ROW)来确定每一景图像的位置。每一景图像实际覆盖地面的面积为185 km×185 km,图1.4所示为轨道120/38采用321波段组合获取的遥感图像。

1972年,英国EMI公司的工程师Outfield发明了用于头颅诊断的X射线计算机断层摄影(Computer Tomograph, CT)装置。这种无损伤诊断技术的基本方法是根据人的头部截面的投影,经计算机处理来重建截面图像。1975年,EMI公司又成功研制出全身用的CT装置,获得了人体各个部位鲜明清晰的断层图像。1979年,这项技术获得了诺贝尔奖。

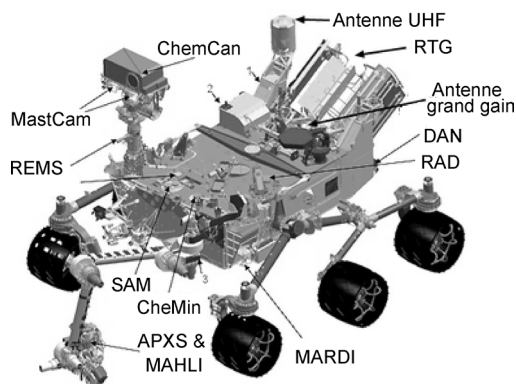


图 1.3 “好奇号”火星车仪器装置

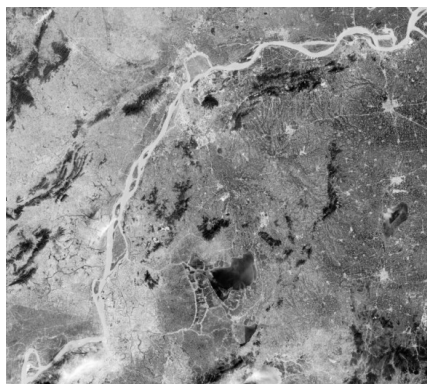


图 1.4 陆地卫星 5 号 TM 图像

20 世纪 70 年代，个人计算机和各种图像输入、输出设备的大众化，为图像处理技术的普及铺平了道路。短短的三四十年，航天、军事、医学和工业等方面应用需求的不断增长，使数字图像处理技术发生了日新月异的变化，完成了从科学研究领域到工程应用的转变，从而进入普及、实用阶段，已经发展成具有强大生命力的学科和产业。特别是 DVD、多媒体计算机、数码摄像机、数码相机、数字电视等信息产品的出现和普及，给图像处理技术带来了无限的生机。数字图像处理理论、技术与设备已经成为当代信息技术重要的组成部分。

从 20 世纪 70 年代中期开始，随着计算机技术和人工智能、思维科学研究的迅速发展，数字图像处理向更高、更深层次发展。人们已开始研究如何用计算机或其他智能装置实现类似人类视觉系统理解外部世界，进入图像理解或计算机视觉阶段。很多国家对这项研究投入了更多的人力、物力和财力，取得了许多重要的研究成果。

在图像处理的理论和方法方面，1965 年快速傅里叶变换 (FFT) 的出现是一个具有代表性的成果，它提供了一种高效率的图像处理工具。在图像处理基础上，人们进一步开展了图像分析和图像理解的研究与应用工作。这些工作需要人类的视觉过程有进一步的理解，甚至需要建立视觉的数学模型，以计算机视觉（机器视觉）来仿真人类视觉。一个具有代表性的成果是 20 世纪 70 年代末美国麻省理工学院 (MIT) 的 Marr 教授提出的视觉计算理论。该理论成为计算机视觉领域其后十多年的主导思想。

随着科学技术的发展，我国在航天、医学、工业、家用信息设备等领域大力研究和应用了图像处理技术。同时，我国已经有一支较强的从事图像工程研究和应用的队伍。1990 年，我国成立了全国性学术团体——“中国图象图形学学会”。它由中国从事图像图形学基础理论与应用研究，软、硬件技术开发及应用推广的专家学者和相关科技工作者组成。国内著名的高等学校、科研院所及 IT 企业都是本学会的重要成员单位。

近年来，图像处理技术为我国的宇航工程立下了汗马功劳。以“嫦娥”登月工程为例，三步探月，“绕”、“落”、“回”，是中国探月工程在 2020 年前设定的三大目标，成像也是探索月球的手段之一。探月工程一期的任务是实现环绕月球探测。“嫦娥一号”卫星所携带的干涉成像光谱仪、激光高度计和 CCD 立体相机共同完成了月球表面三维立体影像的获取工作。2007 年 11 月 20 日开始传回探测数据，经过处理制作完成了第一幅月面图像，并同时完成了三维影像的制作，标志着中国首次月球探测工程取得成功。CCD 相机采用线阵推扫的方式获取图像，轨道高度约 200 km，每一轨的月面幅宽 60 km，图像共由 19 轨图像制作而成，像素分辨率为 120 m。图像覆盖区域属月球高地，主要由斜长岩组成，分布有不同大小、形态、结构和形成年代的撞击坑，右上部暗色区域由玄武岩覆盖。图幅宽约 280 km，长约 460 km。2008 年 11 月 12 日，我国第一幅全月影像图正式发布，它是由“嫦娥一号”卫星拍摄的

589 轨经过各种处理、校正和镶嵌拼接以后的合成照片,成为目前世界上公布的最完整、清晰度最高、层次最分明、位置最精确的月球影像图,其图像质量和影像制作方面均达到国际先进水平。2007 年 12 月,卫星在 100 km 轨道正常运行了 13 天。其间 CCD 立体相机获取了分辨率为 65 m 的月面图像。“嫦娥一号”还首次获得了白天和黑夜的全月球微波图像。2009 年 3 月,在受控撞击过程中,“嫦娥一号”卫星携带的 CCD 相机传回实时图像,其图像清晰。图 1.5 所示为我国第一幅月面图像的飞行效果图。“嫦娥一号”卫星第一幅月面图像的处理过程如图 1.6 所示。

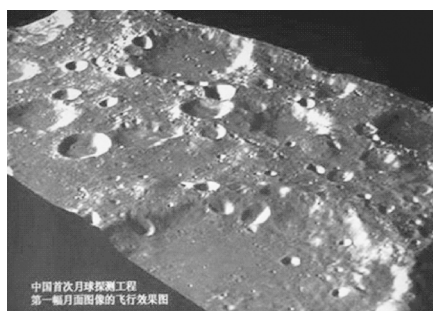


图 1.5 我国第一幅月面图像的飞行效果图

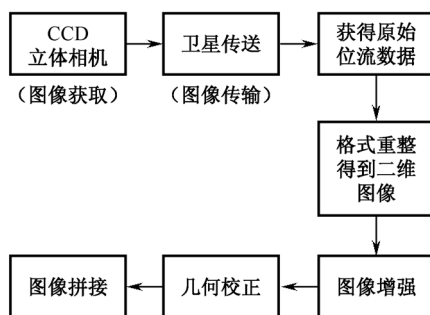


图 1.6 我国第一幅月面图像的处理过程

探月工程二期的任务是实现月面软着陆和自动巡视勘察。“嫦娥二号”的主要任务是为“嫦娥三号”实现月面软着陆开展部分关键技术实验。“嫦娥二号”卫星于 2010 年 10 月 1 日直飞月球,搭载的 CCD 相机在 100 km 圆形轨道和 100 km×15 km 椭圆轨道的近月点处,分别对“嫦娥三号”卫星的预选着陆区进行优于 10 m 和 1.5 m 分辨率的高分辨率成像实验,获得了更清晰、更详细的月球表面影像数据。2010 年 10 月 28 日在虹湾地区,卫星距月面约 18.7 km,分辨率约为 1.3 m。局部三维景观图对应月面东西宽约 8 km、南北长约 15.9 km 的区域。该区域表面较平坦,由玄武岩质的月壤覆盖,分布有不同大小的环形坑和石块,其中最大的环形坑直径约 2 km。“嫦娥二号”卫星的 CCD 相机在 100 km 圆形轨道上的实际分辨率能达到 7 m,进入 100 km×15 km 的椭圆轨道时,其分辨率能达到 1 m。2012 年 2 月国防科工局发布“嫦娥二号”获得的世界首幅 7 m 分辨率全月球影像图。

嫦娥工程第二阶段的“嫦娥三号”卫星实现月球软着陆和自动巡视勘测并开展月基天文观测的登月探测器,包括着陆器和“玉兔号”月球车(巡视器),利用自主研发相机对月球进行近距离的微观小尺度月貌形态进行精细探测成像。2013 年 12 月 2 日,高速摄像机拍摄了发射“嫦娥三号”的长三乙火箭喷着长长火焰的箭体现场,首次应用了箭载的三台摄像机跟踪并实时传回火箭组合体运载火箭发射、飞行状态图像,地面指挥中心首次近距离观察到火箭及卫星组合体的飞行状态。14 日“嫦娥三号”距月球表面 100 m 高度开启变推力发动机,“嫦娥三号”悬停,使用激光测距敏感器和三维成像敏感器获取月面区域的激光三维图像,以躲避大的障碍物为安全软着陆选择合适地点。接着低速向下移动,避开小的月面障碍。当接近 4 m 时,关闭发动机,成功实施自由落体运动软着陆,降落相机传回图像(见图 1.7)。着陆器上携带了地形地貌相机,对着陆器周围进行 360° 成像,绘制着陆区地形、地貌图,用于研究月表的地形地貌特征。地形地貌相机还要面对“玉兔”号进行静态拍照与动态摄影(见图 1.8)。着陆器上还安装了近紫外月基天文望远镜、极紫外相机。巡视器桅杆上装配一套双眼视觉系统的全景相机,利用桅杆 360° 的旋转和 90° 的俯仰获得周边多幅图像,通过图像拼接获得巡视区的全景图像,实现从 3 m 到无穷远的清晰成像。“玉兔”号的导航仪中,导航相机由左右两个相机组成,负责对“玉兔”号行驶路线的偏差进行修正和控制。避障相机也是由左右两个相机组成,固联于车体围栏下方。它还携带了测月雷达,以获取月球内部的物质成分并进行分析,将一期工程的“表面探测”引申至内部探测。巡视器缓缓驶离着陆器后,全景相机就开机工作。15 日,“嫦娥三号”实现

两器分离，并通过安装在着陆器上的地形地貌相机和配备在巡视器上的全景相机互拍成像检验着陆成功，作为“嫦娥三号”任务圆满成功见证。此后，全景相机还肩负着获取巡视器巡视月表区的三维光学图像的科学探测使命。“嫦娥三号”搭载了全新的测控天线、数据传输与处理系统，使用穿透浮尘能力极强的 UHF 频段接收机可以接收并解调巡视器向着陆器发送的遥测数据及科学探测数据，并实时发向地面。



图 1.7 降落相机拍摄视频截图

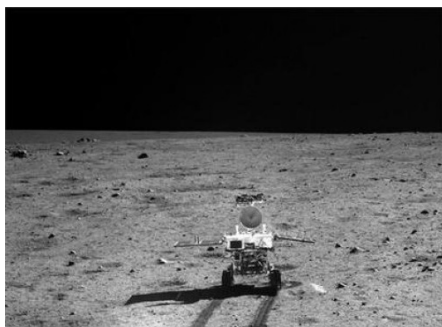
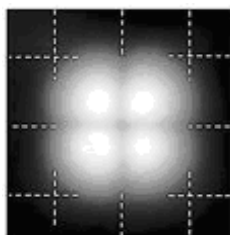


图 1.8 着陆器地形地貌相机拍摄的“玉兔号”巡视图像

依靠显微镜、电子显微镜等设备成像，图像处理同样在微观世界扮演着重要角色。近期我国科学家研制的超高分辨的扫描探针显微镜系统首次拍摄到了清晰的水分子的内部结构图像（见图 1.9），在国际上首次实现了水分子的亚分子级分辨成像，使得在实验中直接解析水的氢键网络构型成为可能。通过解析粉尘表面水结构，有望为解决 PM2.5 危害提供新思路。



(a) 单个分子结构



(b) 4 个水分子组成的水团簇结构

图 1.9 水团簇在盐表面上的氢键网络构型结构

1.3 图像处理的目的是、任务与特点

1.3.1 图像处理的目的是

一般地，图像处理需要完成以下一项或几项任务。

(1) 提高图像的视觉质量以提供人眼主观满意或较满意的效果。图像的增强、图像的恢复、图像的几何变换、图像的代数运算、图像的滤波等处理使受到污染、干扰等因素影响产生的低清晰度、变形图像等的质量得到有效改善。

(2) 提取图像中目标的某些特征，以便于计算机分析或机器人识别。这些处理也可以划归于“图像分析”的范畴。例如，边缘检测、图像分割、纹理分析常用作模式识别、计算机视觉等高级处理的预处理。

(3) 为了存储和传输庞大的图像和视频信息，常常需要对这类数据进行有效的压缩。常用的方法有统计编码、预测编码和正交变换等方法。

(4) 信息的可视化。许多信息（如温度场、流速场、生物组织内部等）并非可视，但转化为视觉

形式后可以充分利用人们对可视模式快速识别的自然能力,更便于观察、分析、研究、理解大规模数据和许多复杂现象。信息可视化结合了科学可视化、人机交互、数据挖掘、图像技术、图形学、认知科学等诸多学科的理论和方法,是研究人、计算机表示的信息以及它们相互影响的技术。

(5) 信息安全的需要。主要反映在数字图像水印和图像信息隐藏方面。这是新世纪图像工程出现的新热点之一。数字水印是利用多媒体数字产品中普遍存在的冗余数据与随机性,把水印信息可见或不可见地嵌入到数字作品中,以期达到保护数字产品的版权或完整性的一种技术。在计算机通信、密码学等学科中,数字水印也有其用武之地。

1.3.2 图像处理的任务

图像处理的任务是获取客观世界的景象并转化为数字图像后,进行增强、复原、重建、变换、编码、压缩、分割等处理,从而将一幅图像转化为另一幅具有新的意义的图像。有时称静止的图像为图片 (picture),而称活动的图像为视频 (video)。

对“图像处理”的理解有广义与狭义之分。广义的“图像处理”,可以包含有些学者提出“图像技术”、“图像工程”等概念,而狭义的“图像处理”重点讨论为改善视觉效果、存储或传输效率,在输入图像和输出图像之间进行的变换。广义的“图像处理”还可以包括“图像分析”(image analysis)和“图像理解”(image understanding)等。图像分析指对图像中感兴趣的目标进行检测和测量,图像理解指在图像分析的基础上,进一步研究图像中各目标的性质和它们之间的关系,以得到对图像反映的场景的合理解释。从抽象程度看,图像处理处于低层,图像分析处于中层,而图像理解处于高层。本书中,“图像处理”指与图像有关的理论、技术和系统,主要介绍经典的图像处理理论和方法,对一些重要的图像分析内容和图像理解的技术热点也做相应的介绍,以扩展读者的知识面,同时进一步感受三者之间的联系与差异。我们可将图像处理的主要任务分成以下几类。

1. 图像获取与数字化

将自然界的图像通过光学系统成像并由电子器件或系统转化为模拟图像信号,再由模拟/数字转换器(ADC)得到原始的数字图像信号。图像的获取也称图像的采集(acquisition)。图像的采集十分重要。原始的图像质量高会大大减轻后期处理的负担。尽管图像处理硬件和软件可以在一定程度上弥补采集过程中存在的缺陷,但保证高信噪比、高保真度的原始图像仍然是首先必须重视的问题。

2. 图像增强

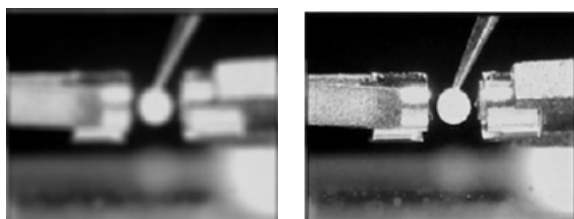
图像增强(image enhancement)的作用是对视觉不满意的图像进行改善,突出图像中所感兴趣的部分。如强化图像高频分量,可使图像中物体的轮廓清晰,细节明显,而强化低频分量可减少图像中噪声的影响,即对高频噪声起平滑作用。可见,尽管人们并不一定知道图像降质(degrade)或退化的原因,但通过使用图像增强技术得到新图像的质量在主观视觉上更为良好。通过图像增强,可以改变原来图像全部或局部的亮度、对比度、色彩分布等参数,使增强后的图像更加赏心悦目。对于图像分析和图像理解来说,图像增强往往作为这些过程的前期处理(预处理),使分析效果更好或更容易理解。图 1.10 针对给定图像中的应用场合,有目的地增强了图像的整体和局部特征,将过暗的部分变得亮度适中,使用去雾(dehazing)功能,使得模糊的原图像变得更加清晰。



图 1.10 图像增强

3. 图像复原

图像复原 (image restoration) 也称图像恢复。如果对图像退化的原因或过程 (如某种噪声的影响、运动造成的模糊、光学系统的几何失真等) 有一定的了解, 通过理论推导或实验数据可以建立退化的数学模型 (“降质模型”), 那么我们可以采用某种滤波方法从降质的图像在一定程度上恢复为原始图像。



(a) 原模糊图像

(b) 逆滤波恢复的图像

图 1.11 逆滤波复原方法

在图像恢复中, 建立图像的退化模型是关键。理论上, 降质的模型一般是非线性、时变和空间变化的, 但这种模型即使使用计算机也很难处理。所以, 在一定的精度下, 用线性、时不变和空间不变的降质模型代替上述模型具有实际意义。图 1.11 是对微装配准备对接模糊图像进行逆滤波方法进行复原的实例。可见恢复后的图像要比原来的模糊图像清晰得多。

图像复原与图像增强都是为了提高图像的质量, 它们之间的区别在于前者需要考虑图像降质的原因, 而后者并不需要这样做。如图 1.10 所示, 去雾的类似效果也可用复原的方法实现。在雾、霾天气下, 大气中的悬浮粒子对于物体表面的反射光有较强的前向散射作用, 使得物体成像的亮度减弱。同时这种散射作用造成了图像饱和度、对比度降低, 以及色调偏移。这就需要建立大气散射模型, 通过分析散射作用对成像的影响, 依据不同方法并借助图像所含信息特征构造场景反照率 (albedo) 和景深 (depth) 的约束条件, 估算出大气传输函数及相关参数进而恢复出清晰图像。

4. 图像重建

图像重建 (image reconstruction) 是一类特殊的图像复原技术。图像的采集是由实际图像产生二维数组 (即二维数据) 的过程。相反的问题是, 如果我们知道一组与图像相关的物理数据, 又如何得到图像呢? 图像重建就是试图从物体横剖面的一组投影数据建立图像。输入是一系列图像, 输出一幅重建图。事实上, 图像重建在医学、工业检测、数据压缩等领域都有重要的应用价值。计算机断层摄影是图像重建的成功范例。物体内部的数据是由各种能量流 (如 X 射线、电子束、超声波等) 穿透物体而获得的。体视学 (volume visualization) 的思想是将图像重建与计算机图形学的光照模型和各种渲染技术相结合, 把多幅二维图像合成为三维图像, 生成高度真实感的图像。体视学在医学领域是应用最早的。现在利用有效的工具可以完成对人体器官、软组织和病变体的三维重建与三维显示。

由于图像是三维景物的二维投影, 一幅图像本身不具备复现三维景物的全部几何信息的能力, 很显然三维景物背后的部分信息在二维图像画面上是反映不出来的。因此, 要分析和理解三维景物, 就必须做合适的假定或附加新的测量, 例如双目图像或多视点图像。在理解三维景物时需要知识导引, 这也是人工智能中正在致力于解决的知识工程问题。

目前主要有两类重建图像的方法: 傅里叶逆变换和级数展开重建技术。通常图像重建的数学模型较复杂, 计算量大, 涉及投影模型、迭代计算等。医学设备中的 CT 和 MRI 都属于投影重建设备。

5. 图像变换

图像阵列很大, 直观性强, 但图像的某些特性 (如频率特性、纹理特性等) 在空间域中难以获得和处理, 计算量也很大。各种图像变换 (image transformation) 的方法, 如离散傅里叶变换 (DFT)、离散沃尔什-哈达码变换 (WHT)、离散余弦变换 (DCT) 和离散小波变换 (DWT) 等, 可以间接地将空间域的处理转换为变换域进行更有效的处理。通过 DFT, 可以将空间域的图像变换为图像频谱,

再在频率域进行各种数字滤波以获得图像质量的改善、数据量的压缩或突出某些后期处理的特征。离散小波变换在空间域和频率域中都具有良好的局部化特性，受到人们的普遍重视。一些图像编码算法已经吸收到国际标准中，如 JPEG 标准采用 DCT 算法，而 JPEG2000 采用 DWT 算法。

6. 图像编码与压缩

图像编码与压缩 (image coding and compression) 在图像存储和传输中起着至关重要的作用。我们知道，数据量庞大是数字图像的显著特点之一。在多媒体技术中，现有的大容量存储器和宽带网络技术仍不能满足对图像数据处理、存储和传输的需要。因此，图像及其他海量数据的压缩是必需的。而且，由于图像等数据中存在相当大的冗余信息，这类数据的无损和有损压缩也是有可能的。编码是压缩技术中最重要的方法，它是发展最早且比较成熟的图像处理技术。

数字图像中相邻像素的相关性较强，说明图像信息压缩的潜力很大。在图像画面上，经常有很多像素有相同或接近的灰度或色彩。就电视画面而言，同一行中相邻两个像素或相邻两行间的像素，其相关系数可达 0.9 以上，而相邻两帧之间的相关性比帧内相关性还要大一些。通过图像编码与压缩，可以实现几倍到十几倍无失真的无损压缩，以及几十倍甚至上百倍允许失真条件下的有损压缩。图像编码压缩技术可减少描述图像的数据量（即比特数），以便节省图像传输、处理的时间并减少所占用的存储器容量。

图像及视频压缩已经渗透到了人们的生活中，VCD、DVD 的普及就是明证。各种图像和视频压缩标准（如 JPEG 标准、MPEG 标准）大力推动了数据压缩技术的进步和相关软件与硬件的产业化。高清晰度数字电视的普及使压缩技术在图像通信领域得到更显著的发挥。尽管未来存储设备的容量会更大、信道带宽会更宽，但数据压缩技术仍然是多媒体系统不可缺少的关键技术之一。

7. 图像分割

图像可以视为由背景和一个或多个目标组成。图像分割 (image segmentation) 是按一定的规则将图像分成若干个有意义或感兴趣区域的过程，每个区域可能代表一个对象（目标或目标的一部分）。通过图像分割，图像中有意义的特征部分（如边缘、区域等）被提取出来。图像的这些特征是进一步进行图像分析和理解的基础。

人眼对图像进行分割比较直观，也很迅速，但由计算机进行图像分割并非易事。我们希望通过分割后的图像更便于计算机或机器人识别和理解，是图像处理向图像分析过渡的一个关键步骤。然而，一般图像的构成是十分复杂的，对图像自动分割十分困难，而分割的结果往往不能令人满意。目前，印刷体的光学字符识别 (Optical Character Recognition, OCR)、指纹识别等应用领域开始使用自动分割技术。大部分图像的分割需要人工干预来提高分割的可靠性和有效性。由于图像分割是一个重要和难度较大的课题，至今仍然是图像处理工作者乐此不疲的研究对象。图 1.12 所示为某汽车车牌识别的系统程序流程图及效果图。

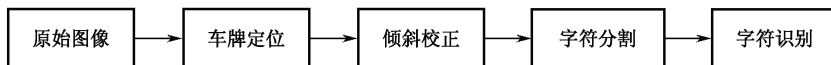


图 1.12 汽车车牌识别的流程图及效果图

虽然目前已研究出不少边缘提取、区域分割的方法，但还没有一种普遍适用于各种图像的有效方法。因此，对图像分割的研究还在不断深入之中，是目前图像处理中研究的热点之一。

8. 图像融合

由于单一图像传感器获取的数据信息量有限，往往难以满足实际需要，而利用多源数据则可以提供对观测目标更加可靠的观察。图像融合 (image fusion) 利用了多源信息进行决策和行动的理论和技

术和工具，将多源信道（传感器、数据库或人为获取的信息）所采集到的关于同一目标的不同成像机理、不同工作波长范围、不同工作环境与要求的图像数据经过图像处理最大限度地提取各自信道中的有利信息，消除多传感器信息之间可能存在的冗余和矛盾，最后综合成高质量的图像，以供观察或进一步处理。图像融合以提高图像信息的利用率、改善图像获取的精度和可靠性、提升原始图像的空间分辨率和光谱分辨率为目的，以形成对目标清晰、完整、准确的信息描述，有利于系统对目标进行可靠的探测、识别、跟踪及情景感知。

各种遥感器所获得的大量光谱遥感图像的分辨率、灰度等级可能相差很大，如能有效融合将为人们提供更加清晰、可用的图像。不同途径的遥感图像首先需要配准，然后再进行融合。2013 年我国公布“高分一号（GF-1）”卫星获取的首批影像图，该卫星是国家重大科技专项——高分辨率对地观测系统的首发星，卫星配置了 2 台分辨率为“2 米全色/8 米多光谱”的高分辨率相机和 4 台分辨率为 16 米的多光谱中分辨率宽幅相机。大同市影像图为“高分一号”卫星首次开机成像获取的图片，体现了卫星多模式同时工作的能力（见图 1.13）。

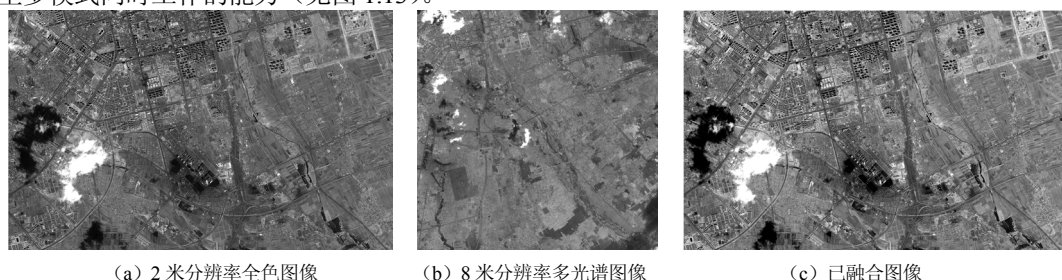


图 1.13 多光谱和全色图像的融合

多传感器图像融合技术最早被应用于遥感图像的分析与处理中，在测绘、地质、农业、气象及军事目标识别等方面得到了广泛应用。到 20 世纪 80 年代末，人们才开始将图像融合技术应用于一般图像处理。随着 20 世纪 90 年代 LANDSAT-7、SPOT-5、RADARSAT、JERS-1、ERS-1 卫星的发射升空，图像融合更成为遥感技术的研究热点。目前图像融合在医学、计算机视觉、对地观测、机场导航、安全监控、智能交通、地理信息系统、工业过程控制、智能机器人等民用领域也得到了广泛应用，甚至成为解决某些难题的关键技术之一。在军事领域，以多传感器（如多热像仪、电视摄像仪及激光测距仪等）技术为核心内容的战场感知已成为现代战争中最具影响力的军事高级技术。在医学领域中，通过多源医学图像融合可以综合不同模态医学图像的优点，从而为医学诊断、人体功能和结构研究提供更充分的信息。CT 和 MRI 图像的融合处理已成功应用于颅脑放射治疗和颅脑手术可视化中。CT 图像与 MRI 图像有很好的信息互补性。CT 的空间分辨率高于 MRI，而 MRI 的对比分辨率高于 CT，特别是软组织对比分辨率明显优于 CT（见图 1.14）。

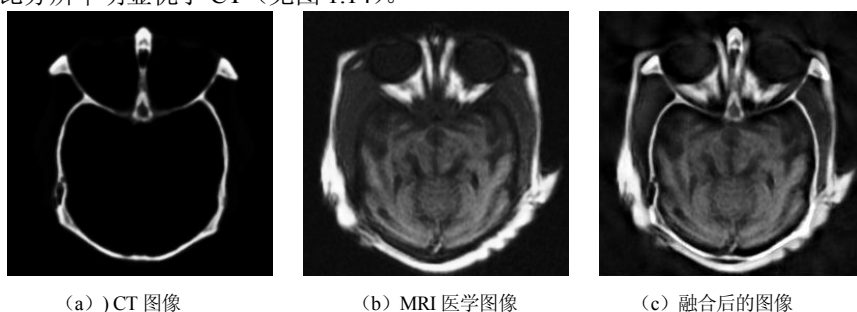


图 1.14 医学图像融合结果

1.3.3 数字图像处理的特点

数字图像处理利用数字计算机或其他专用的数字设备处理图像，与光学等模拟方式相比具有以下鲜明的特点。

1. 具有数字信号处理技术共有的特点

(1) 处理精度高。图像采集设备可将一幅模拟图像数字化为任意大小和精度的二维数组供处理设备加工。根据应用的需求，数字化的像素数可以从几十到几百万，甚至上千万，每个像素的等级可以量化为从1位到16位甚至更高，活动图像的帧率可以从十几Hz到60Hz。而对处理设备来说，不同数据量的图像其处理程序大致是一样的。

(2) 重现性能好。理论上，数字图像处理不会因图像的存储、传输等过程而导致图像质量的退化。图像的质量主要受数字化过程时采样样本数、量化精度、处理过程中的处理精度等的限制。由于在一定范围内，人眼和机器视觉的分辨率都是有限的，所以只要保持足够的处理精度，则数字图像处理过程就能够保证原有图像的重现。

(3) 灵活性高。与模拟图像处理相比较，由于图像处理软件功能十分强大、扩展性好、与用户可以友好地交互，数字图像处理不仅能完成一般的线性和非线性处理，而且一切可以用程序实现的智能信息处理方法都可以加以采用。

2. 数字图像处理后的图像可能是供人观察和评价的，也可能作为机器视觉的预处理结果

如果供人观察，则处理后的图像的质量优劣必然受人的主观因素影响。由于人的视觉系统十分复杂，受环境条件、视觉性能、人的心理和知识背景等因素的影响，其评价体系也难以统一，故通常对图像处理的评价往往从客观和主观两方面进行。另一方面，机器视觉是依靠计算机来模仿人的视觉功能的，通过人类视觉感知机理的研究促进计算机视觉的研究，但图像庞大的信息量、多义性、环境因素的影响以及不同知识的导引，使得机器视觉对图像的理解的正确性远远低于人类视觉。

3. 数字图像处理技术适用面宽

原始模拟图像可以来自多种信息源，它们可以是可见光图像，也可以是不可见的波谱图像，如各种射线图像（见图1.15）、超声波图像或红外图像（见图1.16）。从图像反映的客观实体尺度看，可以小到电子显微镜图像，大到航空照片、遥感图像甚至天文望远镜图像。这些来自不同信息源的图像只要被变换为数字编码形式后，均可用计算机来处理。



图 1.15 胸部 X 光片图像



图 1.16 红外热像仪检测的手

4. 数字图像处理技术综合性强

由于数字图像处理技术适用面宽，涉及的技术领域也十分广泛，从学科分类来看，将其划分为交

叉学科比较合适。数学、物理学（包括光学、电学等）等领域是数字图像处理的理论基础，计算机技术、电子技术、摄影技术、电视技术、通信技术等是其实现的支撑技术。

5. 数字图像处理与模拟方式处理图像相比，也有一些不足之处

（1）数字图像处理的信息大多是二维或二维以上的多维信息，数据量巨大。

一幅中等分辨率的 VGA 640×480×256 色图像的数据量为 300 KB；传送一路 PCM 彩色电视图像的速率达 108 Mb/s，则每秒的数据量可达 13.5 MB。这么大的数据量和传输速度对计算机的计算速度、网络的带宽、媒体的存储容量等提出了很高的要求，因此数据压缩成为不可缺少的处理环节。

（2）数字图像信号占用的频带较宽。

在模拟域，视频信号的带宽比音频信号的带宽约高两个数量级。为了保证图像的质量，根据采样定理，数字化后，数字视频占用的频带进一步加宽。所以，在成像、传输、存储、处理、显示等各个环节的实现上，宽频带对处理和传输设备提出了更高的要求，同时，与数据压缩对应的频带压缩技术也是图像处理的一个值得注意的问题。

（3）处理费时。

由于图像数据量较大，因此处理比较费时。特别是采用区域处理方法时，由于处理结果与中心像素邻域有关而导致花费的时间更多。要实现快速甚至实时处理图像，就要对图像处理系统提出更高的要求，多处理器并行处理器、嵌入式系统等专用处理系统为提高图像处理速度提供了有效的解决方法。

1.4 数字图像处理的应用

图像是人类获取和交换信息的主要来源，图像处理起初主要应用在遥感、医学等领域，然而，随着人类活动范围的不断扩大、需求的不断提高，图像处理的应用几乎渗透到科学研究、工程技术和人类社会生活的各个领域。下面分 6 个方面概要介绍数字图像处理的主要应用。

1. 航天和航空技术方面的应用

上面已经介绍了在航天领域 JPL 采用数字技术对月球、火星照片的开创性处理工作。另一方面，在飞机遥感和卫星遥感技术中，数字图像处理也起到了其他技术无法替代的作用。侦察飞机或卫星获取的大量空中摄影照片需要进行处理和分析，如果人工进行处理和识别，则需要花费大量的人力资源，不仅速度慢，还容易产生人为误差。采用计算机图像处理系统来判读分析，既节省了人力，又加快了速度，还可以从照片中提取人工所不能发现的大量有用情报。从 20 世纪 60 年代末以来，美国及一些国际组织发射了资源遥感卫星（如 LANDSAT 系列）和天空实验室（如 SKYLAB），由于成像条件受飞行器位置、姿态、环境条件等的影响，图像质量不可能很高，必须采用数字图像处理技术进行几何校正、恢复、增强等加工，从而还原图像的本来面目。如 LANDSAT 系列陆地卫星，采用多波段扫描器（MSS），在 900 km 高空对地球每一个地区以 18 天为一个周期进行扫描成像，其图像分辨率大致相当于地面上的十几米或 100 m 左右。飞行器先将成像的模拟图像进行数字化处理，而后将得到的数字信号存入磁带中，在卫星经过地面站上空时，再高速传送下来，然后由处理中心分析判读。现在，通过发射合理分布的卫星星座可以 3~5 天观测地球一次。高分辨率卫星遥感图像可以得到优于 1 m 的空间分辨率。

由飞机遥感和卫星遥感获得的图像数据是极其有用的，利用陆地卫星所获取的图像可以进行资源调查、灾害检测、资源勘查、农业规划、城市规划、气象预报等。20 世纪末，人们提出了“数字地球”的概念，试图通过建立“数字地球”来解决这样一种矛盾，即一方面拥有大量可以发挥积极作用的信息，而另一方面又让它闲置起来慢慢地过时而变得无用。“数字地球”以计算机技术、多媒体技术和大规模存储技术为基础，

以宽带网络为纽带,运用海量地球信息对地球进行多分辨率、多尺度、多时空和多种类的三维描述,从而支持 and 改善人类的活动与生活质量。可见,数字图像处理技术是“数字地球”的重要技术基础之一。

2. 生物医学工程方面的应用

数字图像处理在生物医学工程方面的应用十分广泛,且具有无创伤、快速、直观、准确等优势。数字图像处理在医学上应用最成功的技术要数 X 射线 CT 技术,该技术的主要研制者因此而获得诺贝尔生物学奖。图像处理首先应用于显微图像的处理分析,如红细胞、白细胞分类,染色体分析,癌细胞识别等。在 X 光肺部图像增强、超声波图像处理、心电图分析、立体定向放射治疗等医学诊断方面都广泛地应用图像处理技术。目前,图像存档和传输系统 PACS 的建立将在医学图像处理与分析的基础上为数字医院打下坚实的基础。

3. 通信工程方面的应用

数字图像通信包括传真、电视电话、数字电视、电视会议等。当前通信的主要发展方向是声音、文字、图像和数据结合的多媒体通信,电话网、电视网和计算机网络将以“三网合一”的方式形成多媒体通信网。由于图像的数据量十分巨大,必须采用编码技术来压缩信息的比特量。图像通信特别是高清晰度的视频通信已成为实现多媒体通信的一个瓶颈,编码压缩技术是必须突破的关键技术之一。

第三代(3G)移动通信技术以宽带多媒体为重要特征,作为其终端产品——智能 3G 手机小巧便携,集电话、笔记本电脑和图像通信等于一身,可提供发送彩信、拍摄和传送动态图像、接收标清电视节目等各种信息服务。已经面世的第四代(4G)移动通信技术集 3G 与 WLAN(无线局域网)于一体,并能够传输高质量视频图像。4G 最大的数据传输速率超过 100 b/s,这个速率是 3G 移动通信速率的 50 倍。由于 4G 可以接收高分辨率的电影和电视节目,从而成为合并广播和通信的新基础设施中的一个纽带。

4. 工业自动化和机器人视觉方面的应用

在工业生产领域,图像处理和机器人技术有着广泛的应用,对提高劳动生产率具有重大意义。管理者可以通过监控系统远程监视车间的生产情况,自动装配线中的图像测量装置可以无损检测产品的质量、对产品进行分类。在一些危险、有毒、放射性大、劳动强度大的环境中,利用机器人完成识别工件、装配产品、维护设备更是必不可少的,它们已在太空、深海、重工业、高污染等场合中得到了有效利用。自动识别系统可以实现邮政信件的自动分拣,从而大大减轻邮局工作人员的工作负担。

5. 军事和公安方面的应用

以信息技术为核心的高新技术化已经成为现代战争越来越明显的特征,正在引发一场深刻的军事革命。信息化战争(IW)将逐步取代工业时代的机械化战争,成为未来战争的基本形态。图像信息的获取和利用将在信息化战争中扮演重要角色。图像处理和识别可以用于各种侦察照片的判读、导弹的精确制导。具有图像获取、传输、处理、评估、选择、存储和显示的军事指挥自动化系统(C⁴I, Command, Communication, Control, Computer and Intelligence)是结合以计算机为核心的技术装备与指挥人员,对部队和武器实施指挥与控制的“人-机”系统。同样,为适应未来的信息化战争,坦克、飞机和军舰模拟训练系统等大量采用了包括图像处理技术在内的虚拟现实技术。

图像处理技术在公安业务中也发挥重要作用。特别是生物特征识别技术大量以图像信息(指纹、掌纹、虹膜、面像、步态、笔迹等)作为研究对象,在公安刑侦、出入境管理、智能通关、金融认证、电子商务等领域得到关注。生物特征识别技术也带动了智能卡、条形码等存储生物特征信息技术的发展。图像处理也应用于对各种其他图片的判读分析、不完整图片资料的复制和修复、交通视频监控中的“电子警察”、不停车自动收费系统等方面。

近年来，信息伪装技术发挥着越来越重要的作用。信息伪装将秘密信息隐藏于另一非机密的文件内容之中，其形式可以是任何一种数字媒体，如图像、视频、声音等。在被保护的對象中嵌入某些能够证明版权归属或跟踪侵权行为的信息，这些信息可能是作者的序列号、公司标志、有意义的文本等。水印中的隐藏信息能够抵抗各类攻击。即使水印算法是公开的，攻击者要想毁掉水印仍十分困难。数字水印作为在开放的网络环境下保护版权的新型技术，可以确立版权所有，识别购买者，或者提供关于数字内容的其他附加信息，并将这些信息以人眼不可见的形式嵌入数字图像。在数字视频序列中，用于确认所有权和跟踪行为。此外，数字水印在证据篡改鉴定、数据的分级访问、数据的跟踪和检测、商业和视频广播、互联网数字媒体的服务付费及电子商务的认证鉴定等方面，也具有广阔的应用前景。另一方面，图像真实性鉴定也引起了人们的高度重视，如伪造照片常常引发道德和法律问题。过去需要在暗室里经过几小时繁杂操作才能造假的照片，如今一般人利用计算机软件就可轻易伪造出以假乱真的照片。有些伪造的图像可以通过仔细观察加以判别。然而，在许多情况下，我们需要基于计算机软件工具的数字图像鉴定技术。这些技术需要了解图片的统计特性或几何特性会受到某种具体作弊手法的破坏，然后编制一套算法来找出这些破绽：如从不同的照片上截取所需图像会出现明暗问题：合成一张新照片时，由于拍摄时光线条件不同，不同照片上的人或物会存在细微差别，从而可以发现图像拼合的破绽；检查照片是否存在因缩放操作而产生的像素伪迹；发现 JPEG 压缩存在不一致的漏洞。

6. 材料科学方面的应用

随着材料科学和计算机技术研究的发展，计算机图像分析系统逐渐成为辅助研究材料结构与性能之间定量关系的一种重要的手段。由光学或电子显微镜、光谱等各种材料表征手段可获得有关材料的结构和性能的影像。

以公路交通的沥青为例。沥青是一种复杂的胶体体系。为适应交通荷载和环境的要求，通过对道路石油沥青的改性来提高和改善沥青的各项性能已成为延长沥青路面使用寿命的重要措施，而其中更



图 1.17 SBS 改性沥青材料的显微图像

以聚合物改性沥青的应用最为广泛。在改性沥青的应用过程中，常见的问题是宏观性能相当的改性沥青其使用性能却相差很大，即使是同一种改性沥青其性能亦可相差很远，要解决此类问题必须研究其微观结构。基于显微形态结构的分析方法成为研究改性沥青性能、机理的有效手段。苯乙烯-丁二烯-苯乙烯嵌段共聚物（SBS）改性沥青的原料是基质沥青，利用 SBS 改性剂良好的物理性能对基质沥青做改性处理。通过剪切、搅拌等方法使加入基质沥青中一定比例的 SBS 改

性剂均匀地分散于沥青中，再向材料中加入一定比例的专属稳定剂，共同形成 SBS 共混材料。通过对其显微图像（见图 1.17）的处理和分析，可以获取其数字特征，对 SBS 改性沥青的生产、性能测试和改善具有较显著的意义。

7. 生活和娱乐方面的应用

数字图像处理技术在文化艺术和休闲娱乐方面所起的作用是有目共睹的。PC 处理图像的能力已经达到昔日大型计算机的处理能力。数字摄像机、数码相机、扫描仪、高分辨率打印机等图像输入/输出设备和各种各样的图像处理软件使 PC 如虎添翼。早期的 VCD 和 DVD 的普及，现在的智能手机、平板电脑和数字电视的兴起使数字图像处理设备进入千家万户。这类应用还有电视画面的数字编辑、艺术照片、电子游戏、纺织工艺品设计、服装设计与制作、发型设计、计算机美术、运动员动作分析和评分等。智能手机被设计成为具有较强的图像处理功能，如卫星地图生成、名片识别、二维码识别

等。其中,二维条码是用特定的几何图形按一定规律在平面上分布的条、空相间的图形来记录数据符号信息。通过二维码扫描,用户可以进行解析网址(图 1.18 所示为 QR 码,3 个定位图形位于左上角、左下角、右上角)、购买产品、物流管理、防伪等操作。

第 11 章将详细介绍几种图像处理的具体应用实例及其发展趋势。



图 1.18 微信二维码

本章小结

- (1) 本章主要介绍数字图像处理的发展简史、图像处理的任务、特点及应用。
- (2) 要求掌握图像、数字图像处理、数字图像处理系统的概念及它们之间的相互关系。
- (3) 本章重点要求明确图像处理的目的和任务、数字图像的应用。

思考题与习题

- 1.1 连续图像和数字图像如何相互转换?
- 1.2 采用数字图像处理有何优点?
- 1.3 数字图像处理主要包括哪些研究内容?
- 1.4 说出图像 (image)、视频 (video)、图形 (drawing) 及动画 (animation) 等视觉信息之间的联系和区别。
- 1.5 如何在你的显示器上观测图像的像素? 1 台笔记本电脑的屏幕分辨率为 1366×768 , 长宽比例是 4:3、16:9 还是 16:10? 屏幕分辨率 (以像素每英寸 (DPI) 为单位) 是多少? 选择一个高分辨率图像、一个低分辨率图像, 比较分别观察到像素时图像的放大倍数的差异。
- 1.6 如何在你的计算机上查看图像文件的各种属性?
- 1.7 讨论在你的智能手机中哪些功能应用了图像处理技术。
- 1.8 设图 1.18 的文件名为 “microscopic.jp2” 或 “microscopic.jpg”, 尺寸为 4140×3096 。针对本图像, 讨论可以进行哪些图像处理和图像分析操作?
- 1.9 讨论数字图像处理的主要应用。进一步查找资料, 写一篇你感兴趣的关于应用方面的短文。

第2章 图像处理基本知识

➡ 内容提要

本章介绍人类视觉与色度学基础、图像的表达方法、基本的图像处理系统、图像的基本统计特征、MATLAB 图像处理工具箱、Photoshop 等图像处理软件和开发工具。

➡ 知识要点

- 人类颜色视觉和色度学知识。
- 灰度图像和彩色图像的描述方法，重点讨论灰度静止图像。
- 典型的数字图像处理系统。
- 图像的基本统计特征。
- MATLAB 图像处理工具箱、Photoshop 等图像处理软件入门。

➡ 教学建议

- 本章教学安排 2 学时。
- 学习本章的先修知识主要包括：光度学、矩阵的表示与运算、概率论、数字信号处理、信息论与编码、微机原理、程序设计基础等。
- 重点让读者了解微机图像处理系统、MATLAB 图像处理工具箱、Photoshop 图像处理软件的初步使用。

2.1 人类视觉与色度学基础

人类视觉的产生是一个复杂的过程。除了光源对眼睛的刺激，还需要人脑对光刺激的解释。人感受到的物体颜色主要取决于反射光的特性。如果物体比较均衡地反射各种光谱，则看起来是白的。而如果物体对某些光谱反射得较多，则看起来物体就呈现相对应的颜色。为了进行图像的彩色分析，需建立研究彩色计量的学科，称为色度学（colorimetry）。

2.1.1 人类的基本视觉特性

视觉系统中存在着杆状细胞和锥状细胞两种感光细胞。杆状细胞为暗视器官，锥状细胞是明视器官，在照度足够高时起作用，并能分辨颜色。锥状细胞大致将电磁光谱的可见部分分为三个波段：红、绿、蓝。由于这种原因，这三种颜色被称为三基色，图 2.1 所示为人类视觉系统三类锥状细胞的光谱敏感曲线。

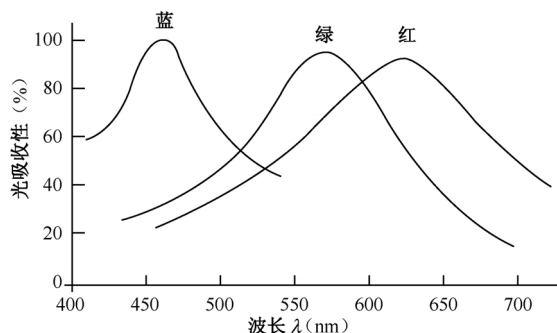


图 2.1 人眼中红、绿、蓝锥状细胞的波长吸收函数

人类视觉对颜色的主观感觉可以直观地用色调（hue）、色饱和度（saturation）和亮度（brightness）来表达。色调表示从一个物体反射过来的或透过物体的光波长，它是由颜色种类来辨别的，如红、橙、绿。色饱和度即色纯度，指颜色的深浅，例如同样是红色，也会因浓度不同而分为深红和浅红。亮度是颜色的明暗程度，从黑到白，主要受光源强弱影响。

2.1.2 三基色原理

由三基色（three primary colors）混配各种颜色通常有两种方法：相加混色法（additive color mixing）和相减混色法（subtractive colour mixing）。彩色电视机上的颜色是通过相加混色产生的。而彩色电影和幻灯片等与绘画原料一样是通过相减混色产生各种颜色的。相加混色和相减混色的主要区别表现在以下三个方面。

（1）相加混色是由发光体发出的光相加而产生的各种颜色，而相减混色是先有白色光，然后从中减去（吸收）某些成分得到各种颜色。当阳光照射到一个物体上时，该物体将吸收一部分具有一定光谱的光线，并将剩下的光线进行反射，反射光线的光谱将决定我们所看见的物体颜色。

（2）相加混色的三基色是红、绿、蓝，而相减混色的三基色是黄、青、品红。也就是说，相加混色的补色就是相减混色的基色。

（3）相加混色和相减混色有不同的规律。

著名的 Grassman（格拉斯曼）定律指出了视觉对颜色的响应取决于红、绿、蓝这三个输入量的代数和。该定律包括如下四项内容：

（1）所有颜色都可以用相互独立的三基色混合得到。

（2）假如三基色的混合比相等，则色调和色饱和度也相等。

（3）任意两种颜色相混合产生的新颜色与采用三基色分别合成这两种颜色的各自成分混合起来得到的结果相等。

（4）混合色的光亮度是原来各分量光亮度的总和。

国际照明委员会（CIE）选择红色（ $\lambda = 700 \text{ nm}$ ）、绿色（ $\lambda = 546.1 \text{ nm}$ ）和蓝色（ $\lambda = 435.8 \text{ nm}$ ）三种单色

光作为表色系统的三基色。这就是 CIE 的 R 、 G 、 B 颜色表示系统。颜色向量 (color vector) C 如图 2.2 所示, 它由红 (Red, R)、绿 (Green, G)、蓝 (Blue, B) 三刺激值所构成的 (R, G, B) 向量的和构成。颜色向量 C 的计算如下:

$$C = RR_0 + GG_0 + BB_0 \quad (2.1)$$

式中, (R_0, G_0, B_0) 称为原刺激值, 是单位向量。称 R 、 G 、 B 为 C 的三刺激值 (tristimulus values)。

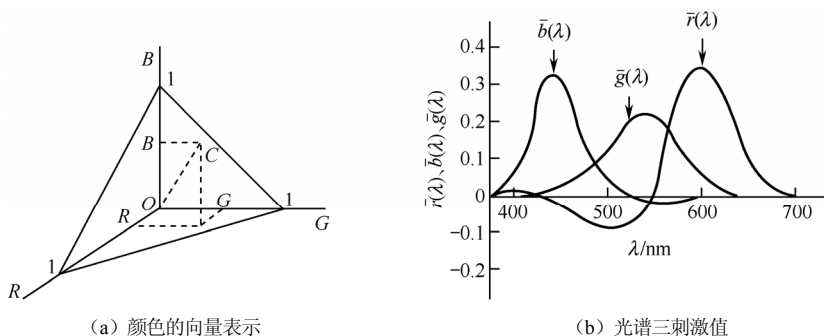


图 2.2 颜色的向量表示与光谱三刺激值

2.1.3 光度学基本知识

光度学是光学中研究光的辐射、吸收、照射、反射、散射、漫射等度量的学科, 同时结合人眼的视觉特征来确定光的度量的单位。同样, 在可见光谱段以外的景物图像也可用类似的方法研究。应该指出, 可见光谱段以外所形成的图像, 其处理的各个过程也常常要变换成人眼可以观察的图像, 例如热成像、X 光照片等都是为便于人对计算机处理过程进行干预, 最终结果映射为可见的图像进行判读。光度学中的几个基本概念简介如下。

1. 光通量 (luminous flux)

光源以电磁波的形式辐射出且人眼所能感受到的光功率称为光通量, 其单位流明 (lm) 的定义为: 发光强度为 1 坎 (cd) 的均匀点光源在一球面立体角内发射的光通量。

2. 发光强度 (luminescent Intensity)

设某个点光源向各方向都均匀辐射, 则可以定义发光强度为发射到单位立体角的光通量增量与该单位角的比值。发光强度的单位为坎 (cd)。

3. 视敏度 (visual acuity)

流明的测量常以人眼的光感觉来度量其辐射功率。人眼对不同波长的可见光的敏感程度不同, 根据人眼对不同波长的敏感程度可得到一曲线, 称为视敏度曲线 (见图 2.3)。

4. 亮度 (luminance, brightness)

亮度是发光面的明亮程度的度量, 它取决于单位面积的发光强度, 单位为 cd/m^2 。

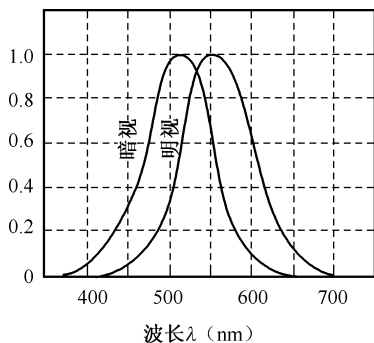


图 2.3 视敏度曲线

5. 照度 (illuminance)

照度指照射在单位面积上的光通量,单位为勒(lx)。图像采集设备的光电传感器对环境光线的灵敏度表示对正常成像时所需要最暗光线的敏感程度。月光级(0.1 lx左右)和星光级(0.01 lx以下)等高增感度摄像机可工作在很暗的条件,红外型可以在没有光线的情况下以黑白图像成像。一般采集设备对中等照度成像效果较好,如学校教室课桌面照度标准值为300 lx。阳光直射的场合照度可高达10万lx,照度过高会导致采集设备进入饱和状态而不能正确成像。

2.2 连续图像的数学描述

人眼所看到的空间某位置上的景物,是由于光照在景物上并经过反射或透射作用映入眼中而形成的图像。因而,一幅图像可以被视为空间各点光强度的集合。我们可以简单地把三维动态图像(3-D image)光强度 I 视为随空间坐标 (x, y, z) 、光线波长 λ 和时间 t 变化的连续函数,典型的表现形式是立体电影和立体电视,其数学表达式为

$$I = f(x, y, z, \lambda, t) \quad (2.2)$$

根据不同波长,图像可分为可见光图像和非可视图像(如X射线图像、红外图像、无线电波图像等)。而遥感图像一般为对同一个场景以多个波段扫描成像的一组图像,即多光谱图像。对于二维图像(如彩色电影和彩色电视)而言,我们可以简单地把光强度 I 视为随空间坐标 (x, y) 、光线波长 λ 和时间 t 变化的连续函数,其数学表达式为

$$I = f(x, y, \lambda, t) \quad (2.3)$$

由于光强度不为负值,所以 $I \geq 0$ 。如果只考虑光的强度而不考虑其波长,则图像在视觉上表现为灰色影像(gray image),称之为动态灰度图像(moving gray image),其图像函数为

$$I = f(x, y, t) \quad (2.4)$$

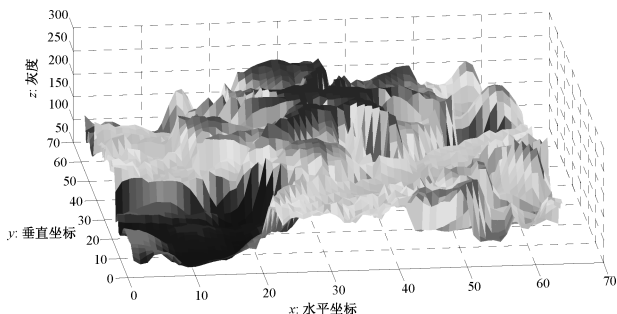
如果处理的灰度图像内容不随时间而变化,则为静止图像(still image),则上述函数可以表示为

$$I = f(x, y) \quad (2.5)$$

图2.4(b)是式(2.5)灰度图像的三维图示,可见垂直坐标 I 与图2.4(a)的灰度是一致的。图2.4(a)对图1.1进行垂直翻转,是为了图(a)和图(b)进行直观的比较。



(1) 图1.1的垂直翻转



(b) 灰度图像的三维图示

图2.4 灰度图像的三维图示

本章重点研究静止图像。理论上,静止彩色图像可以表示为

$$I = f(x, y, \lambda) \quad (2.6)$$

由于彩色图像 (color image) 可以分为红 (R)、绿 (G) 和蓝 (B) 三基色, 因此, 对于彩色图像而言, 图像函数可以用 R、G 和 B 三个通道的值表示:

$$I = \{f_r(x, y), f_g(x, y), f_b(x, y)\} \quad (2.7)$$

可见, 对灰度图像处理的有关理论、方法是开展彩色图像处理工作的基础。在本书第 8 章, 我们也将对彩色图像的处理问题做相应的讨论。

2.3 基本的图像处理系统

数字图像处理系统 (digital image processing system) 是应用计算机或专用数字设备对图像信息进行处理的系统。图像处理系统包括图像处理硬件和图像处理软件。

2.3.1 图像处理硬件

数字图像的显著特点是数据量特别大, 相应地图像操作的时间复杂度也较高, 因此图像处理系统必须配置快速处理硬件, 甚至采用多 CPU 的并行处理技术。早期的数字图像处理系统为了提高处理速度并增加容量, 都采用大型机。以计算机进行图像处理从而改善图像品质的有效应用, 开始于 1964 年美国 JPL 用大型计算机对飞船发回的大批月球照片进行处理, 获得了显著的效果。20 世纪 80 年代后出现了图像处理工作站。对于点处理已采用快速硬件流水线处理器, 它由视频运算器 (ALU) 和查找表 (LUT) 组成, 用于实时进行图像间加、减、乘、除、逻辑运算和灰度变换。对于邻域处理, 应用快速实时小核卷积器, 它由乘法器、累加器、移位寄存器和查找表等组成, 用于实时卷积滤波、去噪声、增强、平滑和边缘提取等。对于大域处理, 可用快速阵列机, 它配有数字信号处理器等用于快速傅里叶变换、各种矩阵运算和向量运算。随着微型计算机性价比的提高, 微型图像处理系统得到迅速发展和普及。主机为 PC, 配以图像采集卡、显示设备和丰富的图像处理软件, 构成了基本的微机图像处理系统 (见图 2.5)。近年来, 为了提高图像采集和处理的实时性, 高端 32 位嵌入式系统应运而生。嵌入式系统是后 PC 时代的主导, 32 位嵌入式系统是电脑硬件与软件的有机结合。嵌入式系统的大体构架可分为五部分: 处理器、内存、输入与输出、操作系统及应用软件。32 位嵌入式系统可分为硬件和软件两个平台。硬件平台的设计包括处理器电路、网络功能、无线通信及使用接口等的设计。嵌入式软件是信息、通信网络等系统中的必备软件, 为硬件产品的驱动程序、控制处理和基本接口功能服务, 以提高硬件产品的价值。嵌入式软件为该硬件产品不可缺少的重要组成部分。下面主要介绍微机图像处理系统。

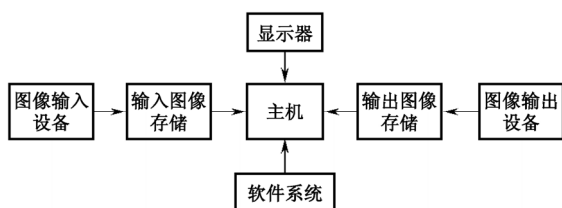


图 2.5 基本的数字图像处理系统

如图 2.5 所示, 微机图像处理硬件系统主要由图像输入设备、图像运算处理设备 (微型计算机)、图像存储器、图像输出设备等组成。软件系统包括操作系统、控制软件及应用软件等。

较完整的数字图像处理系统结构框图如图 2.6 所示。

1. 图像输入设备

图像输入设备完成将模拟光学图像转换成模拟电图像的过程。数字化图像输入设备还进一步将模拟电图像进行数字化, 以便于存储介质存储和计算机处理。根据不同的需求, 图像输入设备主要有电

视摄像机、扫描仪、数码照相机和各种遥感图像获取设备(如红外摄影设备、微波设备)等。第3章将比较详细地介绍主要的图像输入设备的原理及其性能指标。

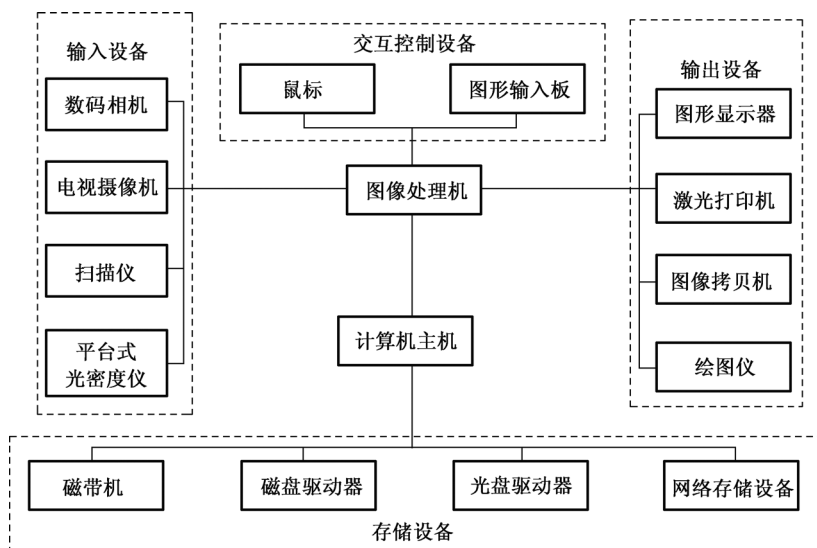


图 2.6 数字图像处理系统结构框图

2. 图像存储器

为了适应图像的大数据量要求,输入图像、输出图像及中间结果图像必须用大容量存储介质进行存储。大容量硬盘、磁带等磁存储设备,CD-ROM、CD-RW 等光学存储装置,以及 SAN (Storage Area Network, 存储区域网)、NAS (Network Attached Storage, 网络附属存储) 等网络存储系统,为海量图像数据存储提供了良好的支持。图 2.7 列举了典型的图像存储器。



(a) 磁带



(b) 数码相机的 SD 卡、CF 卡



(c) NAS/SAN 网络存储设备

图 2.7 典型的图像存储器

3. 主机

主机即用于图像处理的计算机。由于图像数据量大、计算复杂度高,因而对系统硬件配置具有较高的要求。主机的配置主要反映在 CPU、内存和硬盘等指标上。目前主流个人计算机的配置已完全满足一般图像处理的要求。在专业应用场合可以选用图形工作站完成图像处理工作。

为了完成图像处理工作,一些功能卡有时也是不可缺少的。它们包括图像采集卡、显示卡、高速图像处理卡等。

(1) 图像采集卡

如果图像输入装置输出的是模拟信号,则需要在计算机主板扩展槽中安装图像采集卡。采集卡主要包括摄像头 A/D 接口、帧存储器、监视器 D/A 接口和 PC 总线接口单元等。摄像头采集图像数据,经 A/D 变换后将图像存放在图像存储单元中,D/A 变换电路自动将图像显示在监视器上。通过主机发

出指令，将某一帧图像采集到帧存储器中，然后可对图像进行存盘或处理。图 2.8 为典型的图像采集卡外形图。

（2）高速图像处理卡

通常在非实时的场合下，采用软件方案可以实现大部分图像处理的任务。但在工业生产现场、军事等实时应用领域，常常需要用图像处理专用硬件代替图像采集卡，对采集的图像进行实时处理并对其处理结果进行监控。

图 2.9 所示为一种基于 DSP 的高速图像处理卡结构图。通过在图像采集卡上集成高性能的 DSP 芯片而构成高速图像处理卡，由 DSP 替代计算机的 CPU 进行图像处理。摄像机将捕捉到的视频信号输入到 ADC 后，转换成数字视频信号。然后数字视频信号输入到高速 FIFO 中。一旦 FIFO 中的数字视频数据快满时，DSP 将这些数据读入到内部 RAM 中，进行数字视频信号的算法处理。DSP 将最后的运算结果输入给 PCI 总线控制器。PCI 总线控制器以 DMA 方式将运算结果传到主机的内存中。

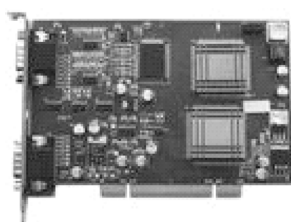


图 2.8 典型的图像采集卡

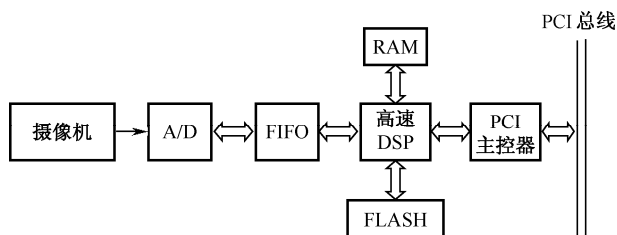


图 2.9 基于 DSP 的高速图像处理卡结构图

（3）显示卡

显示卡简称显卡，是临时保存图像供显示器显像的板卡，安装在计算机主板的扩展槽中。图 2.10 是 VGA 显卡的原理框图。VGA 显卡的分辨率为 640×480 ，16 色。显示内存为具有相同地址空间、同时读写的 4 个 64 KB 的位面，由定序器进行位面选择。图形控制器控制处理器和显示内存间的传输模式，对写入显存的数据进行逻辑操作。CRT 控制器控制对 CRT 的刷新等操作，产生对 CRT 的控制信号。串行发生器将来自显存的数据转换成串行位流。属性控制器将来自串行发生器的颜色值转换成 CRT 要求的模拟信号。定序器控制显卡上各部件的时序操作。

VGA 标准在清晰度和色彩的保真度方面不尽如人意。目前 PC 多采用 SVGA 标准显卡，相应地显示器可显示 800×600 个或 1024×768 个像素点，256 色或 24 位真彩色。其中 24 位真彩色指灰度值红(R)、绿(G)、蓝(B)各占 8 位 (bit)，共计 $2^{24} = 1678$ 万种颜色。

图 2.11 所示为一款 PCIE (PCI-Express) 显卡的外形图。

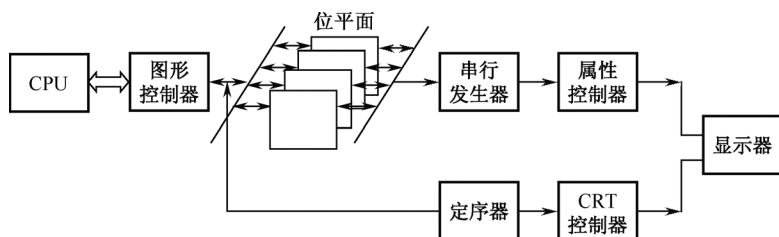


图 2.10 VGA 显卡的原理框图

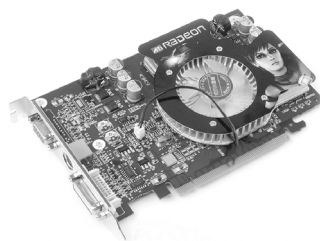


图 2.11 一款 PCIE 显卡

4. 显示器及图像输出设备

显示器是与主机相连接的终端设备，是人与机器沟通的重要界面，同时也是一种必不可少的图像输出设备，通过显示器用户可以随时观察图像处理的中间结果和最终结果，所以有时也将显示器称为监视器。

2.3.2 图像处理软件

图像处理硬件系统也需要软件环境支持才能发挥作用。与各种信息处理系统一样，如图 2.12 所示，图像处理的软件系统也是分层次的。

1. 图像设备驱动程序

上述的各种输入/输出设备、功能卡等硬件装置需要驱动程序才能正常工作。对一般用户来说，主要是安装问题，并不需要自主开发。为了避免用户对硬件了解的困难，Windows 等操作系统提供了“即插即用”（Plug & Play, PnP）功能以简化设备驱动程序开发和设备管理工作，使用户能够方便地添加或删除设备，并允许操作系统在无须用户干预的情况下按照硬件配置的改变而进行调整。即插即用功能包括总线驱动程序、设备安装与配置程序以及从系统注册表中读写信息的自动操作。即插即用提供了对已安装硬件的自动识别、为指定设备驱动程序分配并维护资源、为即插即用系统提供适当接口、为设备事件登记提供相关代码等功能。

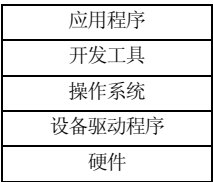


图 2.12 图像处理软件系统的层次结构

2. 操作系统

操作系统是图像处理系统最重要的系统软件，其他软件需要在操作系统的平台上运行。操作系统主要依赖于主机的种类。常见的操作系统有三类：

- （1）IBM PC 及其兼容机一般使用 Microsoft Windows 操作系统。
- （2）Apple Macintosh 机一般使用 Mac OS X 操作系统。
- （3）图形工作站典型使用 UNIX 或 X Windows 操作系统。

3. 图像处理开发工具

图像处理开发工具一般采用图形用户界面（Graphical User Interface, GUI）提供人机交互的工具和方法。用户以某种方式选择或激活这些图形对象（如窗口、按钮、菜单和文本等）时，通常会引起相应的动作。目前图像处理系统开发的主流工具为 Visual C++和 MATLAB 的图像处理工具箱（Image Processing Toolbox，以下也简称为 IPT）。两种开发工具各有所长且相互间有软件接口。

1) Visual C++面向对象可视化集成工具

Microsoft 公司的 Visual C++是一种具有高度综合性能的面向对象可视化集成工具，用它开发出来的 Win 32 程序有着运行速度快、可移植能力强等优点。Visual C++所提供的 Microsoft 基础类库（Microsoft Foundation Class library, MFC）对大部分与用户设计有关的 Win 32 应用程序接口（Application Programming Interface, API）进行了封装，提高了代码的可重用性，大大缩短了应用程序的开发周期，降低了开发成本。

由于图像格式多且复杂，为了使程序员将主要精力放在特定问题的图像处理算法上，Visual C++ 6.0 提供的动态链接库 ImageLoad.dll 支持 BMP、JPG、TIF 等常用 6 种格式的读写功能。该动态链接库共有 22 个函数，其中 18 个函数与图像文件的读写操作有关，这些函数分为三类：Load 函数用于读入图像文件，Save 函数用于保存图像文件，Get 函数用于获取图像文件的相关信息。

OpenCV（Open Source Computer Vision Library）是一个基于开源发行的跨平台计算机视觉库，可以运行在 Linux、Windows 和 Mac OS 操作系统上（见图 2.13）。OpenCV 用 C++语言编写，它的主要接口也是 C++语言，但保留了大量的 C 语言接口，同时提供了 MATLAB 等语言的接口。



图 2.13 OpenCV 的 Logo

OpenCV 实现了图像处理和计算机视觉方面的很多通用算法，致力于真实世界的实时应用，通过优化的 C 语言代码的编写使其执行速度带来了可观的提升，其代码经过适当改写可以运行在 DSP 系统和单片机系统中。

2) MATLAB 的图像处理工具箱和 GUI 开发环境

MATLAB 图像处理工具箱提供了丰富的图像处理函数，灵活运用这些函数可以完成大部分图像处理工作，从而大大节省编写低层算法代码的时间，避免程序设计中的重复劳动。工具箱涵盖了在工程实践中经常遇到的图像处理手段和算法，如图形句柄、图像的表示、图像变换、二维滤波器、图像增强、四叉树分解域边缘检测、二值图像处理、小波分析、分形几何、图形用户界面等。附录 B 列出了常用的 MATLAB 图像处理工具箱函数。MATLAB 还提供了 GUI 开发环境——GUIDE，便于设计者从顶层创建用户图形界面，大大简化了 GUI 设计和生成的过程。设计者可以在 GUIDE 环境下完成输出 GUI 和进行 GUI 编程。第 11 章介绍了基于 MATLAB GUI 图像处理软件的开发过程。

但是，MATLAB 也存在不足之处，这些不足限制了其在图像处理软件中的实际应用。首先，强大的功能只能在安装有 MATLAB 系统的机器上使用图像处理工具箱中的函数或自编的 m 文件来实现。其次，MATLAB 使用行解释方式执行代码，执行速度很慢。第三，MATLAB 擅长矩阵运算，但对于循环处理和图形界面的处理不及 C++ 等语言。为此，通过通用程序接口（API）和编译器与其他高级语言（如 C、C++、Java 等）混合编程将会发挥各种程序设计语言之长，协同完成图像处理任务。API 支持 MATLAB 与外部数据及程序的交互。编译器产生独立于 MATLAB 环境的程序，从而使其他语言的应用程序使用 MATLAB。

3) 图像应用软件

图像应用软件是可直接供用户使用的商品化软件。用户从使用功能出发，只要了解软件的操作方法就可以完成图像处理的任务。对大部分用户来说，商品化的图像应用软件无须用户进行编程，操作方便，功能齐全，已经能满足一般需求，因而得到广泛应用。常用图像处理应用软件有以下几种。

(1) Photoshop

Adobe Systems 公司 1990 年开始推出 Adobe Photoshop，简称 PS，直到 2002 年升级到推出功能十分强大的 Photoshop 7.0 版本，集成了许多成熟并常用的图像处理手段，使用方便，是非常出色的专业图像处理软件，且可以与其他软件交换文件，互相共享数据资源。其他设计软件得到的图像用 Photoshop 再处理，效果更好，处理后的图像又可被其他设计软件、排版软件使用。作为当今世界上一流的图像设计与制作工具，PS 已成为出版界中图像处理的专业标准。2003 年，该公司又推出了 Photoshop 8.0，同时更名为 Adobe Photoshop CS，其中 CS 是 Creative Suite（创造力的工具）的缩写。因此，最新版本 Adobe Photoshop CS6 是 Adobe Photoshop 中的第 13 个主要版本，其典型界面如图 1.14 所示。Adobe Creative Suite 将 Adobe 最新的印刷和 Web 出版工具的完全版本整合在一个功能完善的设计环境中。Adobe Creative Suite 整合了 Photoshop CS 和 Adobe ImageReady CS、Illustrator CS、InDesign CS、GoLive CS、Adobe Acrobat 6.0 Professional 及经过革新的 Version Cue 文件管理器。

高版本的 Photoshop 支持多达 20 多种图像格式和 TWAIN 接口，接受一般扫描仪、数码相机等图像输入设备采集的图像。Photoshop 支持多图层的工作方式，这是 Photoshop 的最大特色。使用图层功能可以很方便地编辑和修改图像，使平面设计充满创意。利用 Photoshop 还可以方便地对图像进行各种平面处理、绘制简单的几何图形、对文字进行艺术加工、进行图像格式和颜色模式的转换、改变图像的尺寸和分辨率、制作网页图像等。

由于 Photoshop 支持多达几十种的文件格式，因此能很好地支持多种应用程序，常见的格式有 PSD、BMP、PDF、JPEG、GIF、TGA、TIFF 等。几种典型格式介绍如下：

① PSD 格式是 Photoshop 的固有格式。图层（layer）在 Photoshop 中是重要的概念，它是构成图

像的重要组成部分, 图层的含义示意图如图 2.15 所示。许多效果可以通过对层的直接操作而得到, 用图层来实现效果是一种直观而简便的方法。使用图层可以在不影响整个图像中大部分元素的情况下处理其中一个元素。改变图层的顺序和属性可以改变图像的最后效果。通过对图层的操作, 使用它的特殊功能可以创建很多复杂的图像效果。PSD 格式可以比其他格式更快速地打开和保存图像, 很好地保存图层、通道、路径、蒙版及压缩方案, 而不会导致数据丢失等。



图 2.14 Adobe Photoshop CS6 的典型界面

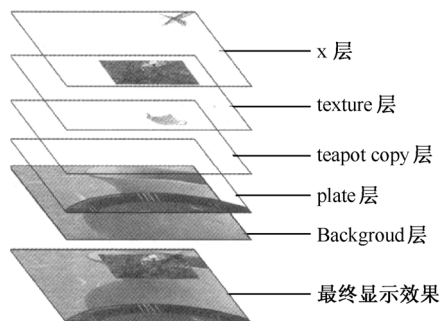







图 2.15 图层的含义示意图

② PDF (Portable Document Format) 格式是由 Adobe Systems 创建的一种文件格式, 允许在屏幕上查看电子文档。PDF 文件还可被嵌入到 Web 的 HTML 文档中。





③ TGA 格式: TGA (Targa) 格式是由美国 Truevision 公司为其显示卡开发的一种图像文件格式, 已被国际上的图形、图像业界所接受, 现在已成为数字化图像及运用光线跟踪算法所产生的高质量图像的常用格式。TGA 文件的扩展名为 .TGA。TGA 的结构比较简单, 属于一种图形、图像数据的通用格式, 目前大部分文件为 24 位或 32 位真彩色图像, 因而能很好地支持真彩色图像, 在多媒体领域有着很大影响。由于 Truevision 公司推出 TGA 的目的是为了采集、输出电视图像, 所以 TGA 文件总是按行存储、按行进行压缩的, 这使得它同时也成为计算机生成图像向电视转换的一种首选格式。

用好 Photoshop, 首先要了解 Photoshop 的界面。如图 2.14 所示, 界面自上而下为菜单栏、反映菜单命令的选项栏, 左边为放置最常用工具的工具栏 (见图 2.16), 右边为各种窗口, 中央为对图像进行处理的图像区域。工具箱中提供了能够辅助我们进行各种操作的有用工具。界面中的一些图形对象 (面板、窗口等) 可以自由组合, 放置非常灵活, 从而方便实现用户的自定义界面。

工具栏中的工具大致可以分为选择工具、矩形工具 (几何绘图工具)、路径工具、文字工具及其他类工具, 此外还有一些提供独立控制功能的按钮和选项。不同的工具使用起来的基本方法都差不多。部分工具图标的小三角标志, 表示该工具是一个工具组, 其中包含有多个与工具组名称相同、相反的工具或同类的工具。下面对常用的工具进行简单的介绍。

① 矩形选框工具 : 用于选择规则形状的工具组, 单击鼠标器右键可以看出该工具组包括矩形选框工具 、椭圆选框工具 、单行选框工具  和单列选框工具 。只要用鼠标拖动图像区域, 用户就可以选择一个矩形或椭圆形区域等。按住 Alt 键的同时单击选框工具可分别切换成矩形选框工具、椭圆选框工具等。使用矩形选框工具可以方便地在图像中制作出长宽随意的矩形选区。如果设置羽化参数, 可以有效地消除选择区域中的硬边界并将它们柔化, 使选择区域的边界产生朦胧渐隐的过渡效果。羽化命令是进行图像合成处理中经常用到的功能, 它能够使我们的图像柔和过渡, 效果看起

来非常自然。矩形选框工具操作时，只要在图像窗口中按下鼠标左键，同时移动鼠标，拖动到合适的大小松开鼠标，即可建立一个简单的矩形选区。图 2.17 所示为矩形选择框选择效果的示例。

② 套索工具：一种用于选择不规则区域形状的工具组。物体的轮廓特征在图像处理中是重要的特征，选择套索工具是为了选取物体的边缘轮廓。为了适应不同的轮廓特征，工具组包括自由套索工具、多边形套索工具、磁性套索工具。使用套索选择工具在图像区域拖动鼠标可以沿着鼠标移动的路径描出所要选取的区域。用套索工具拖动，所到之处留下轨迹。松开鼠标左键，即可形成选择区。磁性套索工具是一种具有自动识别图像边缘功能的套索工具。使用时，将鼠标移动到图像中单击选取起点，然后不用按住鼠标的左键而沿物体的边缘移动鼠标，这时磁性套索工具会根据自动识别的图像边缘生成物体的选区轮廓。当鼠标移回起点时，光标的右下角会出现一个小圆圈，表示选择区域已经封闭，最后在这里单击鼠标即可完成操作。

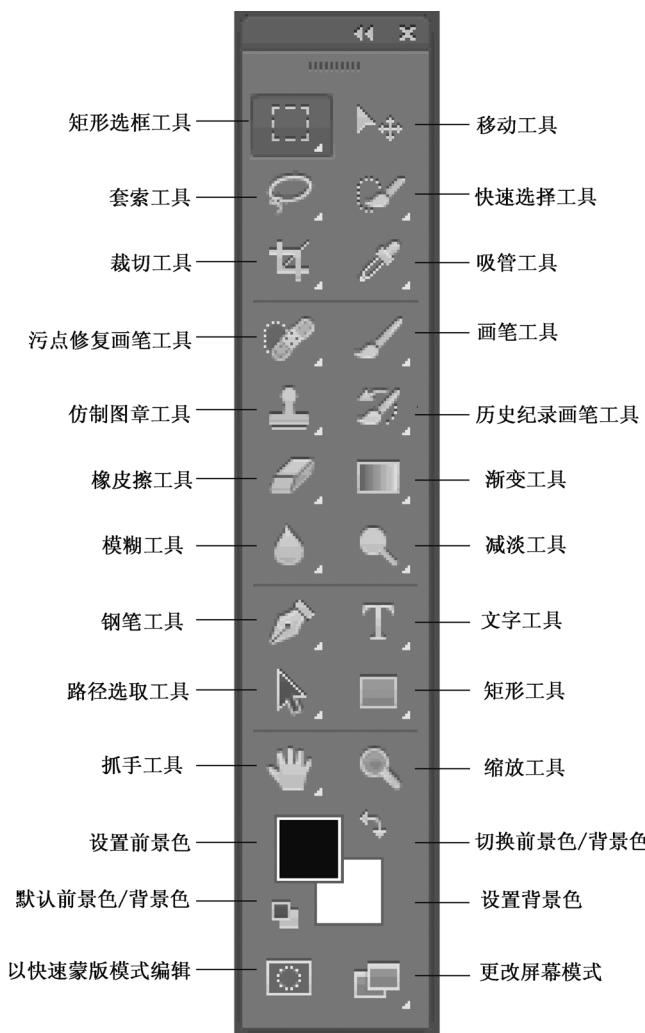








图 2.16 Photoshop 工具栏






③ 快速选择工具：用于选择不规则区域形状的工具组，包括快速选择工具和魔棒工具。魔棒工具是一个有趣的工具，它可以帮助用户方便地制作一些轮廓复杂的选择区。该工具可以把图像中连续或者不连续的颜色相近的区域作为选区的范围。使用魔棒工具只要用鼠标在图像中单击








即可将图像中与光标点中的地方相似的区域变成选择区。魔棒工具的选项栏中包括：选择方式、容差、消除锯齿、连续的、用于所有图层。






④ 移动工具：移动工具可以对被选区域做各种移动。在通常情况下，当被选区域移到原位置后，原来被移掉的部分将以系统的背景填充；当用户按下 Alt 键后再用鼠标拖动被选区域，则原来的背景图像被移走后的部分仍然是原来的图像；当用户按下 Shift 键的同时用鼠标拖动被选区域，则图像只能沿着垂直、水平及 45 度角的方向进行移动。








⑤ 缩放工具：可以方便地改变图像的显示大小，选中缩放工具并单击图像后，图像在屏幕上被放大一级，而按住 Alt 键的同时单击图像将使图像在屏幕上缩小一级。






⑥ 抓手工具：用抓手工具可以在图像窗口不足以显示整个图像时，改变图像显示的位置。





⑦ 画笔工具：常用的绘图工具组，包括画笔工具、铅笔工具、颜色替换工具和工具混合器画笔工具。使用时，只要指定一种前景色，设置好画笔的属性，然后用鼠标在图像上直接描绘即可。在工具的选项栏中，可以选择 Photoshop 中自带的各种形状的画笔，并对它们的各种属性进行设置。





⑧ 矩形工具：方便地制作出各种向量图案的工具组，具有几何绘图的功能，包括矩形工具、圆角矩形工具、椭圆工具、多边形工具、直线工具和自定义形状工具。





⑨ 文字工具：包括横排文字工具、直排文字工具、横排文字蒙版工具和直排文字蒙版工具。在文字工具选项栏中设置好文字的字体和大小等各种属性，随后在图像上需要输入文字的地方单击鼠标左键，开始输入所需的文字，输入完毕，按快捷键 Ctrl + Enter 结束文字编辑状态。

⑩ 吸管工具：吸取选定源图形的属性再将其复制给目标图形工具组，包括吸管工具、3D 材质吸管工具、颜色取样器工具（拾色器）、标尺工具、注释工具和计数工具。吸管工具可以吸取向量线、向量图、位图、文字等的属性。拾色器根据指定位置的色彩值，为调色盘选取所需的颜色。度量工具具有测量工作区域内图像中的任意两点之间距离的作用。

⑪ 前景色和背景色设置工具：前景色指使用绘图工具时的颜色，而背景色指图层的底色，一些用与背景色有关的工具（如橡皮擦、渐变工具等）得到的颜色就是背景色。该工具可以设置前景色（点击左上）和背景色（单击左下），切换前景色和背景色（点击右上），也可以恢复默认的颜色设置（单击左下）。

⑫ 橡皮擦工具：擦除图像的工具组，包括橡皮擦工具、背景橡皮擦工具和魔术橡皮擦工具。在橡皮擦工具的选项栏中可设置画笔的大小与类型。它们的使用方法很简单，像使用画笔一样，先选中橡皮擦工具后，按住鼠标左键在图像上拖动即可。当作用于背景图层时，擦除过的地方会用背景色填充；当作用于普通图层时，擦除过的地方会变得透明。

⑬ 渐变工具：对选区进行填充以营造立体的光影效果的一种绘图类工具组，用于在图像或图层中形成一种色彩渐变的图像效果。渐变工具包括渐变工具、油漆桶工具和 3D 材质拖放工具。渐变工具包括直线渐变工具、辐射渐变工具、角度渐变工具、反射渐变工具、钻石渐变工具。油漆桶工具可以在图像或所选区域中，对指定色差范围内的色彩区域进行色彩或图案填充。

⑭ 减淡工具：用于改变图像的亮度细节、色彩饱和度的工具组。减淡工具包括减淡工具、加深工具和海绵工具。减淡工具和加深工具的原理类似于胶片曝光显影后，通过部分暗化和亮

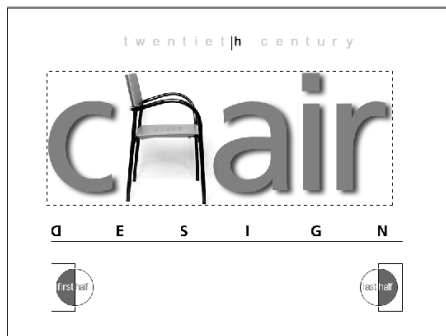






图 2.17 矩形选择框的选择效果

化来改善曝光的效果。海绵工具可以用来调整图像的色彩饱和度,通过提高或降低色彩的饱和度,达到修正图像色彩偏差的目的。

⑮ 模糊工具:一种通过笔刷绘制使图像局部变得模糊、锐化或变形的工具组,包括模糊工具、锐化工具和涂抹工具。模糊工具的工作原理是通过降低像素之间的反差,使图像产生柔化朦胧的效果。锐化工具与模糊工具相反,它可以增强像素间的反差,提高图像的对比度。涂抹工具就好比手指,它可以模仿我们用手指在湿漉漉的图像中涂抹,得到很有趣的变形效果,可以制作出一种类似于水彩画的效果。

(2) CorelDraw

这是 Corel 公司推出的一种基于向量绘图的功能强大的图形图像制作与设计软件。位图式图像是由像素组成的,与其相对,向量式图像以几何、色彩参数描述图像,其内容以线条和色块为主。可见,采用不同的技术手段可以满足用户的设计要求。位图式图像善于表现连续、丰富色调的自然景物,数据量较大;而向量式图像强于表现线条、色块的图案,数据量较小。合理地利用两种不同类型的图像表现方式,往往会收到意想不到的艺术效果。CorelDraw 是当今流行的图像处理软件中为数不多的特点明显、功能强大的基于向量绘图的软件包。利用它,可以方便地制作精美的名片、贺卡、书签、图书封面、广告、宣传画等作品。

(3) ACDSee

这是一个快速、高性能的看图程序,是目前最负盛名的图片浏览器。它能广泛应用于图片的获取、管理、浏览和优化,支持 BMP、GIF、JPG、TGA、TIF、RAW 等超过 50 种常见的图形文件格式,图片打开速度极快,可以直接查看动画 GIF,处理如 mpeg 之类的常用视频文件,还可以为每一个目录建立一个相册。ACDSee 可以从数码相机和扫描仪高效获取图片,并进行便捷的查找、组织和预览。ACDSee 拥有去红眼、剪切图像、锐化、浮雕特效、曝光调整、旋转、镜像等功能,还能进行批量处理,因此是一种极好的图片编辑工具,使用它用户可以轻松地处理数码影像。

各国的科研机构和专业公司还开发了不少专用图像处理软件和应用软件。

2.4 图像的统计特征

图像的种类很多,不同图像的用途和处理方法也不相同。图像处理算法往往针对一定特征(如颜色、纹理、形状、空间关系等)的图像,因此对图像特征的分析至关重要。图像是典型的二维随机信号,其统计特征对于进一步处理和分析是十分重要的。图像的统计特征是图像的一类基本特征。一些算法也是基于图像的统计指标设计的。图像是直观的,但处理的过程常常是抽象的。通过统计分析,可以得到某个图像区别于其他图像的许多参数指标。

由于数字图像为二维矩阵,因此用 MATLAB 处理是十分方便的。基本的图像参数如下。

(1) 图像的大小 (size)

图像的大小是图像的重要属性,其数值是宽度(M,单位为像素)与高度(N,单位为像素)的乘积。调用函数 $S = \text{size}(I)$ 计算图像 I 的大小 $S = [N \ M]$,采用 $K = \text{numel}(I)$ 函数计算图像 I 的像素的总个数 K。

(2) 图像的灰度平均值 (mean)

该指标是指图像 f 所有灰度值的算术平均值 \bar{f} ,反映了一幅图像的平均亮度。

$$\bar{f} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \quad (2.8)$$

调用函数 $\text{ave} = \text{mean2}(I)$ 可计算图像 I 的均值 ave。

(3) 协方差矩阵 (covariance matrix)

设 $f(x, y)$ 和 $g(x, y)$ 是大小为 $M \times N$ 的两幅图像, 则两幅图像之间的协方差矩阵为 C_{fg} :

$$C_{fg} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{f}][g(x, y) - \bar{g}] \quad (2.9)$$

协方差矩阵 C_{fg} 是两幅图像之间的相关程度的一种度量。协方差矩阵为零时表明两图像之间相互独立, 反之表示两幅图像之间相互依赖。 C_{fg} 越大, 相关程度越强。 C_{fg} 是 2×2 矩阵:

$$C_{fg} = \begin{bmatrix} \sigma_f^2 & \text{cov}(f, g) \\ \text{cov}(g, f) & \sigma_g^2 \end{bmatrix} \quad (2.10)$$

式中, σ_f^2 和 σ_g^2 分别是图像 f 和 g 的方差, $\text{cov}(f, g)$ 和 $\text{cov}(g, f)$ 是图像 f 和 g 的协方差。

若 $f = g$, 则单幅图像的协方差即为该图像的方差 (自协方差)。

$$C_{ff} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - \bar{f}]^2 \quad (2.11)$$

调用函数 $C_{fg} = \text{cov}(f, g)$ 计算图像 f 和图像 g 的协方差矩阵 C_{fg} 。调用 $\text{cov}(f)$ 则以每一行为观察样本, 每一列是方差。注意与定义的差异且 f 、 g 均为浮点数 (single 或 double 类型)。将 uint8 类型转换为浮点的函数为 `im2double` 或 `im2double`。

(4) 图像的灰度标准差 (standard deviation)

由于方差按数据与均值差值的平方来计算, 与灰度值相差大, 难以直观地衡量, 所以对方差开平方得到标准差, 标准差是评价离散度的较好指标。

$$SD = \sqrt{C_{ff}} \quad (2.12)$$

调用函数 $SD = \text{std2}(I)$ 计算图像 I 的灰度标准差 SD 。

(5) 图像的相关系数 (correlation coefficient)

协方差可以作为两幅图像相关性的一种度量, 但它受两幅图像像素数值的影响。因此, 经常将其归一化成相关系数。两幅图像 $f(x, y)$ 和 $g(x, y)$ 的相关系数 r_{fg} 是描述图像相关程度的统计量, 是反映两幅图像相互线性联系密切程度的统计参数。相关系数在 $0 \sim 1$ 范围内变动, 其值越接近 1, 两幅图像间的线性相关越密切; 相关系数越接近 0, 相关越不密切。

$$r_{fg} = \frac{C_{fg}}{\sqrt{C_{ff}} \cdot \sqrt{C_{gg}}} \quad (2.13)$$

调用函数 $\text{corrfg} = \text{corr2}(f, g)$ 可以计算大小相等的两幅图像 f 和 g 的相关系数。

【例 2.1】 读取 RGB 彩色图像文件 `football.jpg`, 采用 `rgb2gray` 函数将彩色图像转换为灰度图像。用 `imhist` 计算和显示灰度的统计特性, 求其均值、标准差, 并将图像反白, 另存为一新文件。

【解】 通过如下每行加以注释的程序, 读者可理解求解过程。

```
I = imread('football.jpg'); %用 imread 读取磁盘文件, 并转换成矩阵 I
subplot(2,2,1); %用 subplot 建立多个子图
imshow(I); %用 imshow 显示图像 I
J = rgb2gray(I); %将彩色图像 I 转换为灰度图像 J
subplot(2,2,2); imshow(J); %用 imshow 显示图像 J
subplot(2,2,3); imhist(J); %计算和显示灰度图像 J 的灰度直方图
Ave = mean2(J) %用 mean2 函数求图像 J 的均值
SD = std2(double(J)) %用 std2 函数求图像 J 的均值
s = size(J); %图像大小为 s(1)×s(2)=256×320
```

```

all_white = 255*ones(s(1), s(2)); %设置全部为白色灰度 255
all_white_uint8 = uint8(all_white); %将 double 类型矩阵转化为 uint8 类型矩阵
K = imsubtract(all_white_uint8, J); %图像相减得反白图像 K
subplot(2,2,4); imshow(K); %用 imshow 显示图像 K
imwrite(K, 'football_iverse.jpg'); %用 imwrite 函数将反白后的图像存入文件

```

运行上述程序, 得到 $Ave = 73.9642$, $SD = 37.1148$, $K = all_write - J$, 将图像矩阵存入 football_iverse.jpg, 显示处理结果如图 2.18 所示。

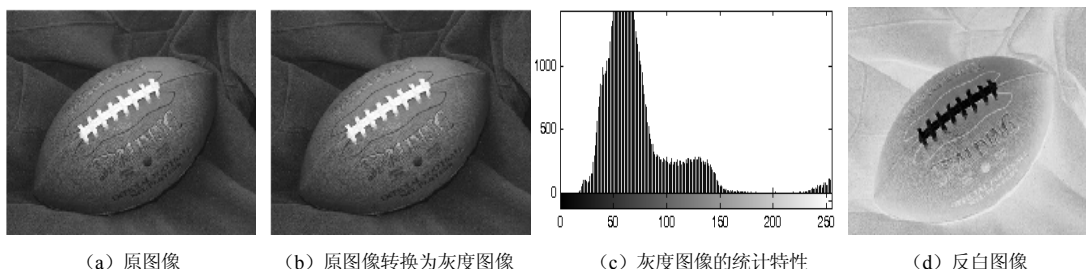


图 2.18 图像统计特性、反白程序运行结果

【例 2.2】 在 PS 环境下, 完成例 2.1 的相关操作。

【解】 (1) 读取图像: 选择“文件”菜单的“打开”命令, 在 MATLAB 安装目录下找到 \toolbox\images\indemos\football.jpg, 即可打开 I 图像并显示在图像区域。

(2) 将彩色图像转换成灰度图像: 选在“图像”菜单的“模式”命令中的“灰度”选项, 即可完成图像颜色属性的转换并在图像区域显示 J 图像。

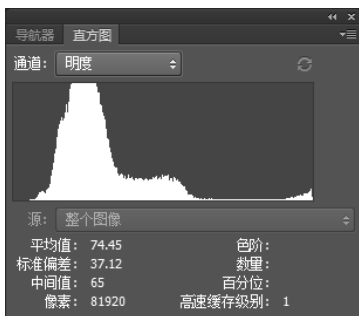


图 2.19 Photoshop 计算的统计数据

(3) 计算 J 图像的统计特性:

图像尺寸: 选择“图像”菜单的“图像大小”命令, 即可查出图像大小为 256×320 。

直方图: 重新打开原文件, 选择“窗口”菜单的“直方图”命令, 即可显示直方图窗口, “通道”选择“明度”, 同时可以查看图像的均值和标准差等统计参数。如图 2.19 所示, 可见与 MATLAB 运行结果是一致的。

(4) J 图像反白操作: 选择“文件”菜单的“调整”命令的“反相”选项, 即可反白 J 图像并在图像区域显示 K 图像。

(5) 保存反白后的图像 K: 选择“文件”菜单的“存储为”命令。

2.5 实验: 图像的基本操作和基本统计指标计算

1. 实验目的

(1) 熟悉 MATLAB 图像处理工具箱和 Photoshop 的基本使用方法。在掌握 MATLAB 基本操作的基础上, 本课程主要依靠图像处理工具箱验证和设计图像处理算法。对于初学者来说, 勤学多练、熟悉 MATLAB 图像处理工具箱和 Photoshop 也是学好本课程的必经之路。

(2) 了解计算图像的统计指标的方法及其在图像处理中的意义。

(3) 了解图像的几何操作, 如改变图像大小、剪切、旋转等。

2. 实验主要仪器设备

- (1) 台式计算机或笔记本电脑。
- (2) MATLAB (安装了图像处理工具箱, 即 Image Processing Toolbox (IPT)) 和 Photoshop。
- (3) 典型的灰度、彩色图像文件、PSD 文件。

3. 实验原理

- (1) 将一幅图像视为一个二维矩阵。
- (2) 利用 MATLAB 图像处理工具箱读、写和显示图像文件。

① 调用 `imread` 函数将图像文件读入图像数组 (矩阵)。例如 “`I = imread('tire.tif');`”。其基本格式为: “`A = imread(filename, fmt)`”, 其中, `A` 为二维数组, `filename` 为文件名, `fmt` 为图像文件格式的扩展名。

② 调用 `imwrite` 函数将图像矩阵写入图像文件。例如 “`imwrite(A, 'test_image.jpg');`”。其基本格式为: “`imwrite(a, filename, fmt)`”。

③ 调用 `imshow` 函数显示图像。例如 “`imshow('tire.tif');`”。其基本格式为: `imshow(I, N)`, 其中, `I` 为图像矩阵, `N` 为显示的灰度级数, 默认时为 256。

- (3) 计算图像的有关统计参数。

4. 实验内容

- (1) 利用 MATLAB 图像处理工具箱和 Photoshop 读、写和显示图像文件。
- (2) 利用 MATLAB 计算图像的有关统计参数。
- (3) 利用 Photoshop 完成选择区域工具、直方图的计算、颜色的调整、图层等操作。

5. 实验步骤

参考附录 B, 常用 MATLAB 图像处理工具箱函数完成以下操作:

- (1) 利用 “读图像文件 I/O” 函数读入图像 `football.jpg`。

(2) 利用 “读图像文件 I/O” 的 `imfinfo` 函数了解图像文件的基本信息: 主要包括 `Filename` (文件名)、`FileModDate` (文件修改时间)、`FileSize` (文件尺寸)、`Format` (文件格式)、`FormatVersion` (格式版本)、`Width` (图像宽度)、`Height` (图像高度)、`BitDepth` (每个像素的位深度)、`ColorType` (彩色类型)、`CodingMethod` (编码方法) 等。

(3) 利用 “像素和统计处理” 函数计算读入图像的二维相关系数 (`corr2` 函数), 确定像素颜色值 (`impxval` 函数), 确定像素的平均值 (`mean2` 函数), 显示像素的信息 (`pixval` 函数)、计算像素的标准偏移 (`std2` 函数) 等。

要求: 参照例题 1.1, 对图像 `J` 加均值为 0、方差为 0.01 的高斯白噪声形成有噪图像 `J1`, 即 “`J1 = imnoise(J, 'gaussian', 0, 0.01);`”, 求 `J1` 的像素总个数、图像灰度的平均值、标准差、`J` 和 `J1` 的互协方差和相关系数、`J` 和 `K` 的互协方差和相关系数。

如果将方差加至 0.1, 重新计算上述统计参数。

(4) 改变图像尺寸 (`imresize` 函数)、旋转图像 (`imrotate` 函数)、对图像进行裁剪 (`imcrop` 函数) 等, 再对操作后的图像进行统计。图 2.20 所示为裁剪后的显微图像。

原图像 `I`, 按比例 `SCALE` 改变尺寸后的图像为 `J`。 `imresize` 函数的调用格式是 “`J = imresize(I,`



图 2.20 裁剪后的显微图像

SCALE)”,同理,图像 A 按 ANGLE 角度进行旋转得到图像 B 的语句为:“B = imrotate(A,ANGLE)”。对于图 1.17 中的 SBS 改性沥青材料的显微图像,由于尺寸偏大,可以按水平和垂直 0.15 的比例减低空间分辨率。对图像 J 进行裁剪的 MATLAB 函数是 `imcrop`,调用格式为“K = imcrop(J, RECT)”,其中 RECT 是 4 元素向量[XMIN YMIN WIDTH HEIGHT],XMIN、YMIN 为左下角坐标值,WIDTH、HEIGHT 分别是裁剪区域的宽度和高度。由于曝光不均匀,可能给图像处理和分析带来困难。舍去暗的地区,采用图像裁剪参数 RECT = [0 181 363 283] 得到光照较均匀的左下角图像,裁剪结果如图 2.20 所示。

要求:参照例 2.1,将图像 I 分别放大和缩小 1.5 倍、旋转 30° ,再对操作后的图像进行统计。

(5) 将经上述不同操作后的图像用“读图像文件 I/O”函数分别写入各自的图像文件。

(6) 在 Photoshop 环境下,查找菜单相应的功能(如读写和显示图像、显示图像文件信息、对图像进行几何操作等),完成上述工作。

(7) 使用选择区域工具。自选合适的图像,使用矩形选框工具、磁性套索工具、魔棒工具对图像进行区域选择并进行适当的处理。后两种工具可以选择非矩形区域。选择所需要的区域,可以针对这些区域进行所需的图像处理。

(8) 选择埃菲尔铁塔、westconcordorthopoto、cameraman、bonemarr、lena 等图像进行直方图的计算。

(9) 选择彩色图像进行颜色的调整。

(10) 图层操作。选择“手.PSD”图像,观测层的有关信息并按层做适当的处理。

6. 实验报告要求

(1) 利用 MATLAB 图像处理工具箱读、写和显示图像文件,计算图像的有关统计参数的原理。

(2) 记录、读取加噪、几何变换处理后的图像及图像的统计数据。

(3) 列出上述图像处理的程序。

(4) 回答思考题。

(5) 撰写心得和体会。

7. 预习要求

(1) 了解 MATLAB 图像处理工具箱和 Photoshop 的基本功能。

(2) 了解图像基本指标特征的含义。

8. 思考题

(1) 说明图像的统计特征对图像处理的意义。

(2) 比较不同图像的统计特征的差别,并分析原因。

(3) 讨论施加不同方差的噪声,对图像统计参数产生的影响。

(4) 图像旋转后,图像的尺寸是否发生变化?如果数据增加,多余的数据是怎样补充的?对图像的统计参数会产生怎样的影响?

(5) 在 Photoshop 中,比较矩形选择区域工具、套索工具和魔棒工具的特点。

(6) 如何利用调整层的操作实现既保留结果又保留原来的文件?

本章小结

(1) 本章主要介绍了数字图像处理的基本知识。从人类视觉特性和色度学基本原理出发,完成对图像进行数学描述,介绍了基本的图像处理系统、图像开发环境及常用的图像处理软件。

(2) 学完本章, 要求掌握色度学基本概念, 了解常用数字图像处理系统的组成原理, 逐步熟悉常见图像处理软件的操作方法。

(3) 本章重点要求掌握图像基本统计特征的意义和计算方法。

思考题与习题

2.1 如何表示图像中一点的彩色值? 颜色模型起什么作用?

2.2 色调、色饱和度和亮度的定义是什么? 在表征图像中一点的颜色时, 各起什么作用?

2.3 按照波段的成像图像讨论图像的分类。

2.4 设有大小为 32×32 的图标, 图标的每个像素有 8 种颜色, 共有多少种不同的图标? 如果每 100 万个可能的图标中有一个有意义, 识别一个有意义的图标需要 0.1 s, 则选出所有有意义的图标需要多长时间?

2.5 讨论数字图像处理系统的组成。列举你熟悉的图像处理系统并分析它们的组成和功能。

2.6 常见的数字图像处理开发工具有哪些? 各有什么特点?

2.7 常见的数字图像应用软件有哪些? 各有什么特点?

2.8 讨论数字图像处理的主要应用。进一步查找资料, 写一篇你感兴趣的关于应用方面的短文。

2.9 调用函数 $C_{fg} = \text{cov}(f, g)$ 计算出图像 f 和图像 g 的协方差矩阵是一个 2×2 的矩阵。说明该矩阵各元素的含义, 以及数值大小对图像特性的影响。

第3章 图像的数字化与显示

➡ 内容提要

本章主要介绍图像数字化的方法，详细讲解图像量化中的标量量化和向量量化方法，同时介绍常用的数字化图像输入/输出设备，为正确选择和使用图像处理的终端设备做必要的知识准备。

➡ 知识要点

- 数字图像的获取方法，重点讲解图像的采样与量化。
- 标量量化和向量量化。
- 图像的输入/输出设备。

➡ 教学建议

- 本章建议安排4学时。
- 先修知识包括数字电子技术基础、信号与系统，以及数字信号处理中有关模拟信号数字化等。
- 要求理解图像的描述方法，重点掌握二维采样和量化的概念，了解常用输入/输出设备的原理和性能指标。

3.1 图像数字化的基本过程

3.1.1 图像的采样和量化

一般地,一个完整的图像处理系统输入和显示的都是便于人眼观察的模拟图像。图像的成像是个复杂的过程。以数码相机为例,为了拍摄高质量的照片,必须正确调节焦距、曝光量(主要由快门速度和光圈控制)、图像分辨率、画质等级等参数,其中数字图像采集(image acquisition)过程要完成模/数转换(A/D)。如图3.1所示,图像数字化处理所需的主要硬件如下。

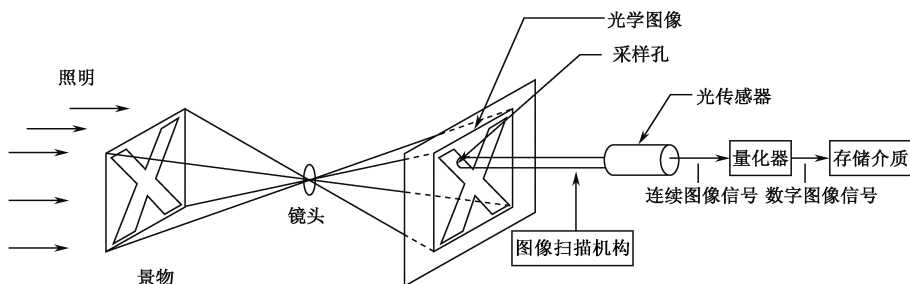


图 3.1 图像数字化硬件示意图

- (1) 采样孔 (sampling aperture): 保证单独观测特定的像素而不受其他部分的影响。孔径越小,成像后的图像空间分辨率将越高,如高档数码相机可高达 2100 万 (5616×3744) 像素。
- (2) 图像扫描机构 (scanning mechanism): 使采样孔按预先确定的方式(一般是从上到下、从左到右)在成像面上移动,从而将二维空间的图像信号转换为一维时间的图像信号。
- (3) 光传感器 (light sensor): 起到光电转换的作用,通过采样孔测量图像的每一个像素的亮度并转化为模拟电信号。目前,图像传感器一般为 CCD 传感器和 CMOS 传感器。
- (4) 量化器 (quantizer): 将传感器输出的连续量转化为整数值,其精度以 bit 计,如 8 bit、10 bit、12 bit、14 bit,甚至 16 bit 等。
- (5) 输出存储体 (output storage media): 将量化后的数字图像存入固态存储盘 (Solid State Disk) 或磁盘等存储介质中。基于闪存的 SSD 常用 Flash 芯片作为存储介质,内部没有机械结构,因此没有数据查找时间、延迟时间和寻道时间。可以移动,数据保护不受电源控制,能适应于各种环境。

现在的单反相机记录的 RAW 原始数据记录文件是感光元件记录光子数量等的数据集。从数码底片 RAW 数据文件通过机内软件(或者后期用户通过计算机的一些软件)转到可视图像文件(如 JPEG 文件)的过程中,色温(白平衡)、亮度、对比度、感光度、色彩等诸多参数的调整因素都会影响到最终输出的画质。高端相机的 RAW 记录文件还能选择 12 bit 和 14 bit。

在一些不适合于人工检测和作业的场所,常用机器视觉(machine vision)来替代人工视觉,以显著提高生产效率和集成化、自动化制造程度。典型的工业机器视觉系统对核心部件的要求很高,这些部件包括:光源、镜头、相机、图像采集卡、机器视觉(或图像处理和分析)软件、监视器、通信/输入/输出单元等。近年来,机器视觉的应用十分迅速,主要在目标检测和机器视觉两个方面。

图 3.2 所示为像素和灰度级的概念。可见,在成像过程中把一幅图像分割成一个个称为像素(pixel)的小区域,这时连续图像就变成了离散图像,而各小区域的灰度用整数型的灰度级(gray level)来表示,这样离散图像就进一步转换成了数字图像。图 3.2 (a) 显示了一个方形像素,图 3.2 (b) 显示了 8 bit 分为 256 个离散级别的灰度级,其范围从黑(0)到白(255)。每一采样的连续灰度值量化为一个离散灰度级。

可见，图像经采样和量化后的结果是一个整数矩阵。假如一幅连续图像 $f(x, y)$ 被采样，则产生的数字图像有 M 行和 N 列。坐标 (x, y) 的值变成离散值，采用整数表示。这样，原点的坐标值是 $(x, y) = (0, 0)$ 。沿图像第 i 行的第 j 个坐标值用 $(x, y) = (i, j)$ 来表示。注意这里 x 和 y 表示高度（垂直）方向和宽度（水平）方向的坐标。图 3.3 表示了 3 行 4 列图像的坐标约定。

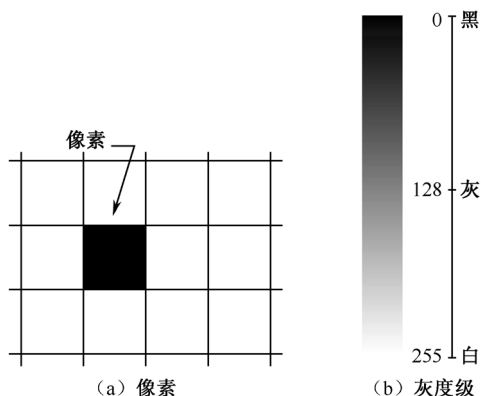


图 3.2 图像的数字化结果

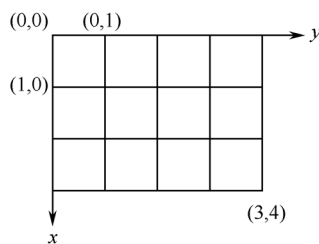


图 3.3 图像的坐标约定

以矩阵的形式表示图像既直观又利于图像处理运算。一幅行数为 M 、列数为 N 的图像可以表示成大小为 $M \times N$ 的矩阵形式 f ：

$$f = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (3.1)$$

式中，矩阵中的每个元素代表一个像素。注意：MATLAB 将图像定义为二维数组，坐标约定从 $(1,1)$ 开始。

【例 3.1】用 MATLAB 程序说明实际数字图像的量化特点。采用 cameraman 测试图像，计算图像的大小，取中央窗口的 16×16 子图像，显示其量化值。对测试图像进行适当裁剪并显示处理后的图像。

【解】实现上述要求的程序 MATLAB 程序如下。

```
W = 16;      %窗口的宽度
I = imread('cameraman.tif');
S = size(I);
J = I(S(1)/2-W/2:S(1)/2+W/2-1,S(2)/2-W/2:S(2)/2+W/2-1)
%取图像中央的子图像，大小为 W*W
imshow(J);
K = I(2*W:S(1)-W,2*W:S(2)-5*W); %裁剪：上 2W, 下 W, 左 2W, 右 5W
imshow(K);
```

运行结果如图 3.4 所示。

图 3.4 (a) 是测试图像，它已是连续图像经过采样和量化后的数字图像，大小为 256×256 。图 3.4 (b) 取中央窗口的子图像（高度和宽度均已放大 16 倍），其量化值如图 3.4 (c) 所示，均为正整数。图 3.4 (d) 是对图像进行适当裁剪后的图像。



图 3.4 实际图像的采样与量化

3.1.2 数字图像的数据量

假定图像尺寸为 $M \times N$ ，每个像素所具有的离散灰度级数为 G ，这些量分别取为 2 的整数幂 m 、 n 和 k ，即 $M=2^m$ ， $N=2^n$ ， $G=2^k$ ，那么存储这幅图像所需的位 (bit) 数是

$$b = M \cdot N \cdot G = 2^{m+n+k} \quad (3.2a)$$

相对于图像高度 M 和图像宽度 N ，存储每个像素所用的位数即为图像深度 G 。字节数为 $b/8$ （单位为 B 或 byte）。如果图像是正方形的，即 $N = M$ ，则所需的位数是

$$b = N^2 \cdot G = 2^{2m+k} \quad (3.2b)$$

可以看到存储一幅图像所需的存储空间随尺寸的增加显著增长。例如，存储上述的一幅大小为 256×256 、灰度级为 256 的图像，需要 $256 \times 256 \times 8 = 524\,288 \text{ bit} = 64 \text{ KB}$ 的空间。

一般分辨率的图像像素数为 512×512 和 256×256 ；高分辨率的图像像素数为 $2\,048 \times 2\,048$ 、 $1\,024 \times 1\,024$ 等。灰度级一般用 8 bit 表示，高精度用 12 bit 或 16 bit 表示。

3.1.3 采样和量化参数的选择

数字图像采样和量化参数的选择直接影响图像的数据量，实际上也会影响到人们的视觉效果和对图像的进一步处理。理论上，采样点和量化等级越多，图像质量越高。但在超出视觉辨别和机器识别的需求时，过多的采样点和量化等级对提高图像质量没有实际意义。对一幅图像，当量化级数 G 一定时，采样点数 $M \times N$ 对图像的空间分辨率有着显著的影响。如图 3.5 所示，当采样点数减少时，图上的块状效应（马赛克效应）就逐渐明显。另一方面，如图 3.6 所示，当图像的采样点数一定时，量化级数越少，图像质量越差。只有两个量化等级的二值图像是量化级数最小的极端情况，如图 3.6 (f) 所示。

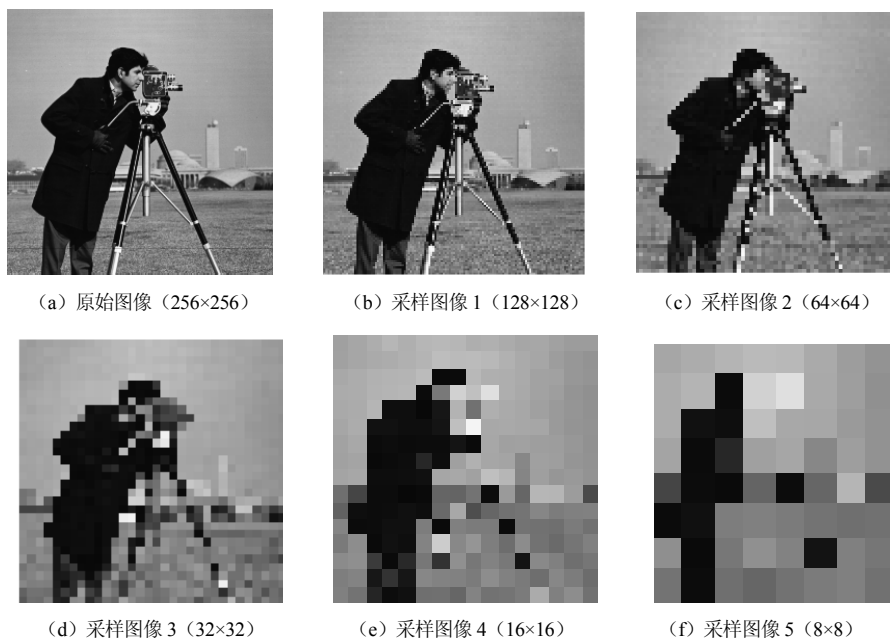


图 3.5 不同采样点数对图像质量的影响

一般，当限定数字图像的大小时，为了得到质量较好的图像，可采用如下原则：

- (1) 对缓变的图像，应该细量化、粗采样，以避免出现假轮廓。
- (2) 对细节丰富的图像，应细采样、粗量化，以避免模糊。

CT 和 MRI 等成像技术一般是在 512×512 像素的采样点、12 位灰度级下对断层扫描图像信息进行数字化采集的，而数字化 X 光片的容量达到 $2048 \times 2048 \times 12$ 位。



图 3.6 不同量化级别对图像质量的影响

3.1.4 二维采样定理

图像在采样时, 应当按照二维采样定理的要求, 无失真或有限失真地恢复原图像。在一维采样定理的基础上, 不难直接导出二维采样定理。

定义二维图像信号 $f(x, y)$ 的傅里叶频谱为 $F(u, v)$, 二维傅里叶变换为

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (3.3)$$

其逆变换为

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv \quad (3.4)$$

二维 Nyquist (奈奎斯特) 采样定理: 如果二维图像信号 $f(x, y)$ 的傅里叶频谱 $F(u, v)$ (下一章将专门介绍) 满足

$$F(u, v) = \begin{cases} F(u, v), & |u| < U_c, |v| \leq V_c \\ 0, & |u| > U_c, |v| > V_c \end{cases} \quad (3.5)$$

式中, U_c 和 V_c 分别对应于空间位移变量 x 和 y 的最高截止频率, 则当采样周期 Δx 和 Δy 满足

$$\left. \begin{aligned} \frac{1}{\Delta x} = u_s \geq 2U_c \\ \frac{1}{\Delta y} = v_s \geq 2V_c \end{aligned} \right\} \quad (3.6)$$

式中, u_s 和 v_s 为采样频率, 它们分别与 Δx 和 Δy 成反比关系。满足上述“等于”条件的采样称为 Nyquist 采样, 满足“大于”条件的采样称为“过采样”, 否则称为“欠采样”。此时, 通过采样信号 $f(m\Delta x, n\Delta y)$ (m, n 为整数) 能唯一地恢复原图像信号 $f(x, y)$, 且有

$$f(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m\Delta x, n\Delta y) \cdot \frac{\sin \frac{\pi}{\Delta x}(x - m\Delta x)}{\frac{\pi}{\Delta x}(x - m\Delta x)} \cdot \frac{\sin \frac{\pi}{\Delta y}(y - n\Delta y)}{\frac{\pi}{\Delta y}(y - n\Delta y)} \quad (3.7)$$

及

$$F_p(u, v) = \Delta x \Delta y \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m\Delta x, n\Delta y) e^{-j2\pi(m\Delta x u + n\Delta y v)} \quad (3.8)$$

一个二维采样信号的频谱示例如图 3.7 所示。图 3.7 (a) 是原图像 $f(x, y)$ 之频谱 $F(u, v)$ 。图 3.7 (b) 是采样后离散图像的频谱 $F_p(u, v)$, 它是 $f(x, y)$ 之频谱的周期性延拓。与一维采样定理类似, 如果满足采样定理的条件, 则各周期频谱彼此分开, 可以用一个合适的二维低通滤波器滤除其他周期频谱成分, 由原频谱来恢复原始信号。否则欠采样将会使各周期频谱成分彼此混叠, 不能由重建滤波器恢复原图像。如果原图像信号的最高频率超过截止频率, 则要求设计抗混叠滤波器来过滤截止频率以上的信号

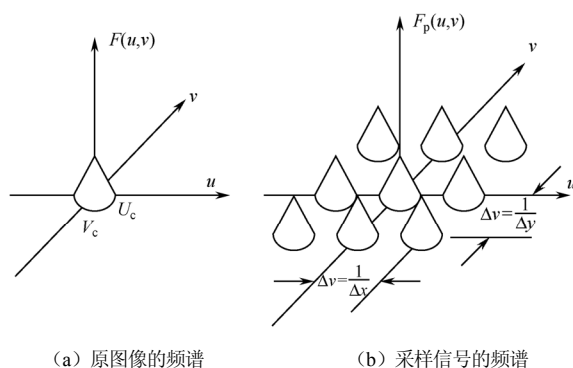


图 3.7 采样信号的频谱

分量。在实际的采样系统中,采样脉冲宽度效应、光学系统的透镜散焦、孔阑衍射等会导致图像模糊,但也会起到低通滤波的作用。

注意,采样后实际图像的大小为有限值,则图像可表示为 $f(m, n)$ ($m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1$)。

3.2 图像的量化方法

3.2.1 量化方法分类

可以从不同的角度将量化方法分成4类。

(1) 按量化级步长的均匀性,可分为均匀量化和非均匀量化。均匀量化是最常采用的量化方式。当图像的灰度级数必须较小时,常常需要非均匀地分配灰度级。例如,由于人眼对于灰度级的急剧变化比较容易辨认,因此这时没有必要过细地分层;而对于那些灰度级变化比较平缓的部分,就需要较细微地分层。否则,不仅无法辨认灰度的细微变化,而且由于较大的量化误差,可能会形成虚假的细节。

(2) 按量化对称性可分为对称量化和非对称量化。对称量化是最常用的量化模式。

(3) 按量化时采样点相互间的相关性,可分为无记忆量化和有记忆量化。前者假定各采样点独立量化,是最常采用的量化方式。后者要考虑先前采样点的影响。

(4) 按量化时处理的采样点数,可以分为标量量化和向量量化。标量量化(scalar quantization)中的“标量”是针对向量量化中的“向量”而言的,标量量化就是将数值逐个量化,而向量量化(Vector Quantization, VQ)则是把一个以上的数值分成一组,组成一个向量,然后按组进行量化编码,如两个数一组或者 $4 \times 4 = 16$ 个数一组一起进行量化。

3.2.2 标量量化

基本标量量化的特点是每次只量化一个采样,而且按照同样的量化准则,前后采样间的量化互相独立。

设在量化前的取值范围为 $[z_0, z_k)$, 概率分布为 $p(z)$, 量化将 $[z_0, z_k)$ 分成 k 个子区间: $[z_0, z_1)$, $[z_1, z_2), \dots, [z_i, z_{i+1}), \dots, [z_{k-1}, z_k)$, 并用一个量化值 q_i 来表示属于子空间 $[z_i, z_{i+1})$ 的数值。

量化器设计完成的任务就是划分子区间和设定量化值,即选定 $k-1$ 个 $z_i (i=1, \dots, k-1)$ 和 k 个 $q_i (i=0, \dots, k-1)$ 使量化造成的失真最小,达到最优量化。失真的度量有两种,一种是客观的度量,使 k 个子区间的总误差平方

$$\varepsilon^2 = \sum_{i=0}^{k-1} \int_{z_i}^{z_{i+1}} (z - q_i)^2 p(z) dz \quad (3.9)$$

为最小时失真最小;另一种是主观的度量,当人眼看不出造成的失真时,失真最小,这时误差小于视觉阈值。这里仅讨论子区间总数 k 已定的情形,若 k 增加,失真当然会减小,但此时所需编码的总比特数也增大了。

1. 均匀量化

最简单的量化方案是均匀量化,又称为线性量化。即把样本值的整个取值范围均匀地分成 k 个子区间。均匀量化将 $[z_0, z_k)$ 均分成 k 个子区间后,每个区间的长度为

$$L = (z_k - z_0) / k \quad (3.10)$$

各子区间以它的中心位置

$$q_i = (z_i + z_{i+1}) / 2 \quad (3.11)$$

作为量化值。当待量化值在 $[z_0, z_k]$ 区间内均匀分布时，也就是说 $p(z)$ 等于常数时，均匀量化是最优的，此时 ε^2 最小，即

$$p(z) = 1/(kL) \quad (3.12)$$

$$\varepsilon^2 = L^2 / 12 \quad (3.13)$$

2. Max 量化器

Max 量化器是一种非均匀量化器，它的设计思想是基于信息论的观点，目标是使 ε^2 最小。其主要思想为： $p(z)$ 不等于常数，说明样本值在某个取值范围内较频繁出现，而在另外一些范围内出现不多。因此可对样本值较频繁出现的取值范围采用较小的量化区间，量化得精细一点，而在其他地方用较大的量化区间，量化得粗糙一些，这样就可在不增加量化级数 k 的条件下，降低平均误差，减少量化噪声。

设 $z \in [z_0, z_k]$ ，量化阶数为 k ，判决电平为 z_i ($i = 0, 1, \dots, k-1$)， q_i 表示量化电平。那么当输入信号的的范围是 $z_i \leq z \leq z_{i+1}$ 时，选择均方差量化器进行量化的误差表示为

$$\varepsilon^2 = \sum_{i=0}^{k-1} E[(z - q_i)^2] = \sum_{i=0}^{k-1} \int_{z_i}^{z_{i+1}} (z - q_i)^2 p(z) dz \quad (3.14)$$

为使误差 ε^2 最小，令其对 z_i 和 q_i 的偏导数为0，可以解出 z_i 和 q_i 。

$$z_i = \frac{q_i + q_{i+1}}{2} \quad (3.15a)$$

$$q_i = \frac{\int_{z_i}^{z_{i+1}} z \cdot p(z) dz}{\int_{z_i}^{z_{i+1}} p(z) dz} \quad (3.15b)$$

Max 量化器在总误差平方和最小的意义上是最优的。但是应用在图像处理中，也有它的缺陷。一般而言，图像在0附近出现的概率较高，因此Max 量化器在0附近必然量化间隔很密，量化较精细。而实际中，人眼在0附近的分辨率并不灵敏，所以用Max 量化器量化得太细是没有意义的。

3.2.3 向量量化

向量量化是一种高效的信源编码技术，它将一组采样的信号幅度向量在容许的误差范围内用更少的离散向量代替。由于它能有效地利用量化的数据采样值间的相关性，向量量化相对于标量量化而言，具有更多的优点。与标量量化相比，当重构层数给定时，向量量化提供较低的失真；或者当失真给定时，提供的重构层数较少。但是由于向量量化编码器需要在码书中进行搜索，找出最优匹配位置，运算量比标量量化大得多。因此，直到20世纪80年代向量量化才开始用于图像编码。也正因为其实现的困难性，目前未找到它与视觉匹配的好方法，向量量化最终未被选入国际标准。尽管如此，仍有不少系统采用了向量量化或DCT和向量量化的复量化方法。向量量化的专用芯片也已出现。

向量量化一次要量化两个以上的采样点。图3.8所示为一个一维的例子。例子中，一次量化4个采样，即一次量化一个四维向量。其量化过程需要用到一个码书，假设我们已经训练好一个含4个码字的码书，每个码字都是四维向量。向量量化过程的实质就是在码书中找到输入向量 \mathbf{X} 的最近码字，其衡量标准就是误差测度。通常人们采用平方误差测度，将向量 \mathbf{X} 与第 i 个码字 \mathbf{C}_i ($i = 1, 2, 3, 4$)之

间的误差记为 $\varepsilon = d(X, C_i)$ 。经过计算发现码字 C_2 离输入向量 X 最近，故用码字 C_2 量化输入向量 X 。最终用码字索引 2 对输入向量 X 进行编码。所以向量量化过程可用式子 $\arg \min[d(X, C_i)]$ 来表示， \arg 表示求索引。

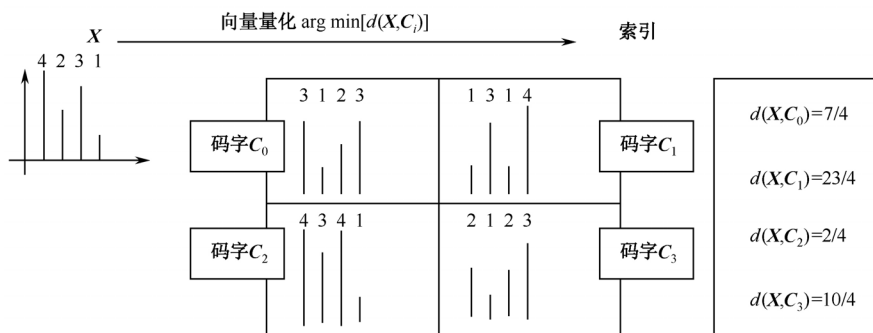


图 3.8 一维向量量化

再看一个图像编码的例子。这里，原始图像块是一个 4 灰度级的 16 维向量 [1 0 0 0 1 1 0 0 2 2 1 0 3 2 1 1]，如图 3.9 所示。向量的每个分量就是一个像素的灰度值，其灰度有 4 个等级，0 最黑，3 最亮。假设码书含有 4 个 16 维码字，如图 3.10 所示。经均方误差准则计算可以发现码字 Y_1 离 X 最近，故用索引 01 进行编码。

 Y_1  Y_2  Y_3  Y_4

图 3.9 原始图像和灰度级

图 3.10 码书 $C = \{Y_1, Y_2, Y_3, Y_4\}$

向量量化中码书的设计按照一定的准则，如最佳匹配、最小失真等。向量量化的过程是一个模式匹配的过程，有多种方法实现这一匹配过程。常用的方法是编码器建立从码书输入码向量的索引。解码器采用标号从查找表中获得重建向量输出，如图 3.11 所示。

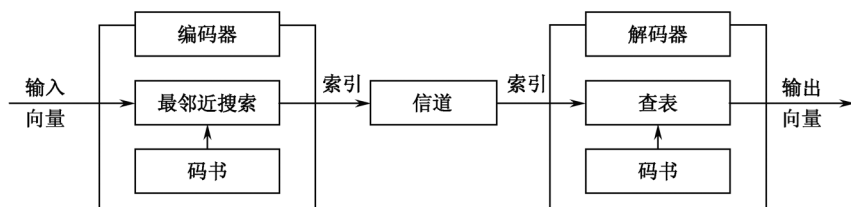


图 3.11 向量编码示意图

3.3 图像输入/输出设备

3.3.1 图像输入设备

数字化器 (digitizer) 是将模拟图像转换成数字图像的数字化输入装置。图像数字化器的性能指标主要包括：像素大小、图像大小、被测对象的局部特征、线性度、灰度级和噪声等。数字化输入设备

的类型很多,常用的数字化器有摄像机、数码相机、扫描仪、数字化仪、CT机等。图3.12是一些典型的数字化设备。



图3.12 图像数字化设备

1. 数码摄像机

数码摄像机(digital video camera, 俗称DV)将图像采集和数字化部件集成,使输出的信号能直接为计算机所接收。数码摄像机的通配性能好,携带方便,适用于现场数据采集。利用感光二极管(photodiode)进行光电转换,将图像转换为数字数据,用于获取活动图像(即视频)信号,是一类首先被广泛使用的图像获取设备,其光电转换器件主要是固态光电转换器件。固态摄像机的传感器类型有电荷耦合器件(Charge Coupled Device, CCD)、COMS图像传感器等。CCD摄像机采用CCD器件替代摄像管实行光电转换、电荷存储与电荷转移。CCD是大规模集成电路(VLSI)的产品,由于体积小、重量轻、寿命长、成像速度快、成本低、灵敏度高、图像无几何失真,已成为视频信号的常用输入设备。CCD摄像机拍摄的图像质量与CCD的数量、CCD的感光面积、CCD的工作方式有很大关系。按CCD的数量可分为:单片、三片式摄像机,三片式摄像机的质量最好。专业机型往往采用三片式,红绿蓝三基色光由3块不同的感光芯片来处理,效果比单芯片的摄像机的成像性能明显要好。

在一些低端的摄像机中,也广泛采用CMOS传感器。一般地,CMOS传感器在灵敏度、分辨率、随机噪声、固定图像噪声、色彩均匀性等成像质量指标方面低于CCD传感器,但是CMOS传感器在单一电源、低功耗、集成化、多功能化、低成本、感光度和抗光晕等方面具有明显优势,现在内部处理电路做得也越来越好,成像质量方面已经能够满足日常拍摄的需要,耗电方面却要比CCD小很多倍。究竟是CCD好还是CMOS图像传感器好,这常是用户困惑的问题。其实两者各有所长。随着CCD和CMOS技术的进步,它们之间的差异一般已经无法区分。

灵敏度、分辨率、信噪比是摄像机的三大指标。

(1) 灵敏度:通常用最低环境照度要求来表明摄像机灵敏度,数字CCD摄像机的灵敏度可达1~3 lx,月光型、星光型和红外型会更低。黑白摄像机的灵敏度大约是0.02~0.5 lx,彩色摄像机多在1 lx以上。0.1 lx的摄像机用于普通的监视场合;在夜间使用或环境光线较弱时,可选择0.01 lx的摄像机。与近红外灯配合使用时,也必须使用低照度的摄像机。电子快门(Electronic Shutter, ES)的时间为1/50~1/100000秒,自动电子快门方式可根据环境的亮暗自动调节快门时间以得到清晰的图像。

(2) 分辨率:用“水平像素数×垂直像素数”表示,其数值越大成像越清晰。图像质量有标清(640×480,约31万像素;720×576,约41万像素等)到高清HDV(720p:1280×720,约92万像素;1080i:1920×1080,约196万像素)等,总像素从80万到400万。分辨率往往是用户最关心的指标,但其上限受到像素极限尺寸、CCD制造工艺、设备性价比等因素限制,并不能无限提高。人们正在研究采用信号处理的超分辨率技术力求解决这些问题。

(3) 信噪比:当摄像机摄取较暗的场景时,监视器显示的画面就比较昏暗,如果信噪比较低,此时容易看到画面中雪花状的干扰噪点。DV的信噪比的典型值一般为45~55 dB。采用12 bit ADC和高精度数字处理技术可使信噪比大于63 dB。

2. 数码相机

数码照相机（Digital Camera, DC）是传统照相技术与数字图像处理技术相结合的产物，主流的感光器件是 CCD（电荷耦合器件）和 CMOS 阵列，通过模数转换（A/D）可将摄取的图像以数字格式（通常为 JPEG 格式）存放起来。感光器件的尺寸是成像质量的重要参数，一般从 1/1.6 至 1/2.5 英寸，越大成像越清晰，全画幅单反（Single Lens Reflex, SLR）达 36×24 mm。数码相机摄取图像的像素数一般在 500 万以上。去马赛克效应、去模糊、图像压缩等处理采用 32 bit RISC 或 DSP 芯片完成，处理后的图像在闪存（flash memory）卡中存储，或到 LCD 显示屏显示，或通过 USB 接口下载到计算机存储以进一步进行处理。部分型号的数码相机可以直接通过彩色打印机打印，具有可进行短暂录像的视频输出口。其原理结构如图 3.13 所示。

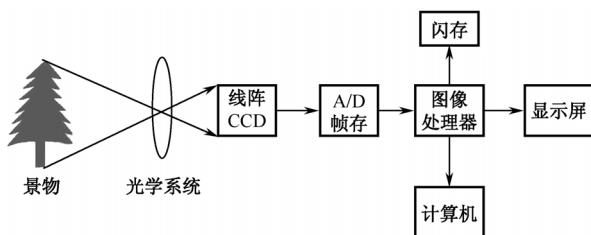


图 3.13 数码相机的原理结构图

数码相机是由光学镜头、CCD 或 CMOS 器件、A/D、MPU（微处理器）、内置存储器、LCD（液晶显示器）、PC 卡（可移动存储器）和接口（计算机接口、电视机接口）等部分组成的。它是一种能够进行拍摄，并通过内部处理把拍摄到的景物转换成以数字格式存放图像的特殊相机。光学镜头的主要参数如下：① 焦距（ f ）：决定视场角的大小。焦距数值越大，视场角越小，便于观测远距离的目标。② 光圈系数（ F ）：反映镜头的通光量，以焦距 f 和通光孔径 D 表示。通光量与 F 值的平方成反比，而镜头的快门速度与其通光量成反比。在不同的场景下，应选取合适的光圈成像。③ 景深（depth of field）：对某景物调焦，以形成前后清晰区域。景深与 3 个因素有关：光圈、焦距和物距。④ 镜头的分辨率：描述镜头的光学传递函数与畸变。而用户关心的是空间分辨率，以每 mm 可以分辨的黑白条纹数测定。⑤ 镜头接口。镜头和摄像头之间的接口有许多不同的类型。C 接口和 CS 接口是工业相机最常见的国际标准接口。与普通相机不同，传统相机利用的是胶片，而数码相机主要靠感光芯片及记忆卡。数码相机可以直接连接到计算机、电视机或打印机上。高档数码相机配置适合影像传输的专用型数字化高清晰度多媒体接口即 HDMI（High Definition Multimedia Interface, HDMI），其最高数据传输速度为 5 b/s，所具备的额外空间可应用在日后升级的音视频格式中。一个 1080p 的视频和一个 8 声道的音频信号的需求少于 4 b/s，这允许它可以用一个电缆分别连接 DVD 播放器等设备。

根据传统照相技术的要求，数码相机的分辨率、感光度和动态范围对于照片成像质量至关重要，因此 DC 应该在不同档次上突出某些指标的优势并平衡上述要素，以满足不同的需求。例如，第一代商业化超级 CCD 很好地平衡了这些要求；第二代超级 CCD 进一步提高了分辨率；第三代超级 CCD 大大提高了感光度；第四代超级 CCD HR 飞跃性地提高了分辨率，同时，超级 CCD SR 大大拓宽了数码相机输出照片的动态范围；第四代超级 CCD 对提高数码相机的成像品质具有划时代的意义。现在高档 DC 像素达 1600 万以上，ADC 精度达 12 或 14 bit，支持高清电影（HD movie）或全高清电影（full HD movie）。

CCD 是亮度敏感器件，不能识别颜色。为此，数码照相机用红、绿和蓝三个彩色滤镜，将输入图像分解成红、绿、蓝三个基色的图像，这样就可以得到每种基色的亮度。通过软件可以对基色图像数据进行处理，从而确定每一个像素点的颜色。

数码单镜头反光照相机（Digital Single Lens Reflex camera, DSLR camera），简称单反相机，是 DC 的一个重要分支，其优秀的成像质量是很多消费者青睐 DSLR 的第一理由。“单镜头”是指摄影曝光

光路和取景光路共用一个镜头, 不像旁轴相机或者双反相机那样取景光路有独立镜头。“反光”是指相机内一块平面反光镜将两个光路分开: 取景时反光镜落下, 将镜头的光线反射到五棱镜, 再到取景窗; 拍摄时反光镜快速抬起, 光线可以照射到感光元件上。因为 DSLR 中感光器的面积远大于一般消费级相机中感光器的面积, 所以像素密度相对大大降低, 因此在宽容度、解像力和高感光度下的表现甚佳。实际上, 感光器的尺寸指标对成像的重要程度远在像素数量之上。消费级相机的感光器尺寸最大不过是 1/1.7 英寸, 而单反的感光器面积则大多是 APS-C 规格 (Advanced Photo System type-C, 23.7mm×15.6mm) 直到与 135 底片一样大小的全幅面尺寸 (36mm×24mm)。

3. 扫描仪

扫描仪 (scanner) 是将各种形式的图像信息输入计算机的重要工具, 也是应用最为广泛的图像数字化设备, 其特点是图像光学分辨率高、设备的成本低, 缺点是获取图像信息速度慢, 不能实现实时输入。

按扫描原理, 扫描仪分为以 CCD 为核心的平板式扫描仪、手持式扫描仪, 以及以光电倍增管为核心的滚筒式扫描仪; 按色彩方式, 分为灰度扫描仪和彩色扫描仪; 按扫描图稿的介质, 可分为反射式 (纸质材料) 扫描仪、投射式 (胶片) 扫描仪, 以及既可扫描反射稿又可扫描投射稿的多用途扫描仪。

一般采用线阵 CCD 器件作为光电转换器件, 其原理结构如图 3.14 所示。扫描仪内部具有一套光电转换系统, 通过光电转换, 将各种图片信息转换成图像数据, 并传送给计算机, 再由计算机进行图像处理、编辑、存储、打印输出或传送给其他设备。其工作过程如下: 首先扫描仪的光源发出均匀光线照到图像表面, 再经过 A/D 转换, 把对应扫描线的图像转换成电平信号, 同时步进电机驱动扫描头移动, 读取下一次图像数据, 经过扫描仪的 CPU 处理后, 图像数据暂存到缓冲器中, 为输入计算机做好准备, 最后按照先后顺序把图像数据传输至计算机并存储起来。

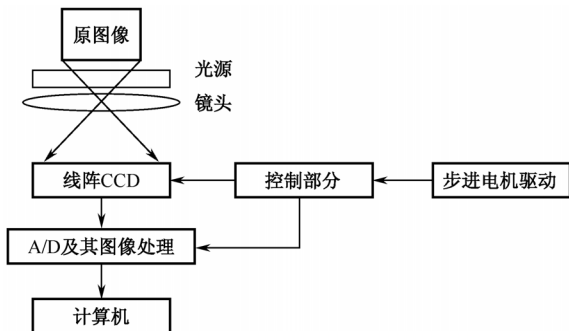


图 3.14 CCD 扫描仪的原理结构图

扫描仪的技术指标主要有: 幅面大小、分辨率、颜色数、接口方式等。幅面大小有 A4、A3、A1、A0 几种, 以 A4 最为常见。

分辨率有光学分辨率和最大分辨率之分。前者指扫描仪光学系统物理器件实际达到的真实分辨率, 后者指经过插值运算获得的分辨率。由于最大分辨率相当于插值分辨率, 并不代表扫描仪的真实分辨率, 所以我们在选购扫描仪时应以光学分辨率为准。分辨率越高, 图像越细腻, 同时数据量越大, 处理的复杂度也将越高。扫描仪的光学分辨率通常用两个数字相乘, 如 600×1200 线, 其中前一个数字代表扫描仪的横向分辨率, 后一数字则代表纵向分辨率, 即扫描仪所用步进电机的机械分辨率。扫描仪的步进电机的精度应该与扫描仪的横向分辨率相同, 但由于各种机械因素的影响, 步进电机的精度将远远达不到横向分辨率的水平。扫描仪的纵向分辨率一般是横向分辨率的 2 倍, 有时甚至是 4 倍。

颜色以位数表示, 有 36 位彩色、42 位彩色、12 位灰度、14 位灰度等。接口方法指扫描仪与主机的接口标准, 有 SCSI、EPP、USB 等。不同的接口方式反映了不同的扫描速度。SCSI 是高速接口, 新型的增强并行端口 (Enhanced Parallel Port, EPP) 已经达到 SCSI 接口的速度水平, 新一代的扫描仪一般采用速度较快的 USB 接口。

手持式扫描仪体积较小、重量轻、携带方便, 但扫描精度低, 扫描质量较差。平板式扫描仪是市

场上的主力军，主要产品为 A3 和 A4 幅面扫描仪，其中又以 A4 幅面的扫描仪用途最广、功能最强、种类最多，分辨率通常为 600~1 200 dpi，高的可达 2 400 dpi，色彩数一般为 30 位，高的可达 36 位。滚筒式扫描仪一般用于大幅面图像扫描，如大幅面工程图纸的输入，它通过滚筒带动图像旋转和扫描头相对位移实现扫描。

扫描仪一般都配有相应的软件，这些软件可用来选择扫描时的工作参数，如扫描区域、对比度、分辨率、图像深度等。此外，有些扫描仪的配套软件还具有平滑、放大、缩小、旋转、编辑等功能。

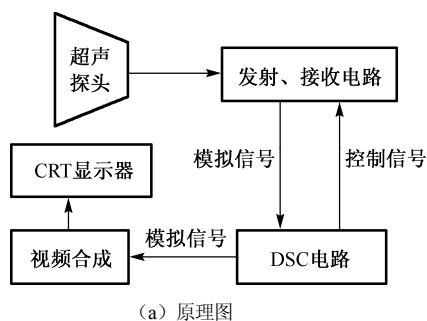
数字化器的性能可以参见表 3.1。

表 3.1 图像数字化器的性能评价

性 能	含 义
空间分辨率	单位尺寸能够采样的像素数。由采样孔径与间距的大小和可变范围决定
灰（色）度分辨率	量化的等级数（位深度）、颜色数（色深度）
图像大小	仪器允许扫描的最大图幅
量测特征	数字化器所测量和量化的实际物理参数及精度
扫描速度	采样数据的传输速度
噪声	数字化器的噪声水平（应当使噪声小于图像内的反差）
其他	黑白/彩色、价格、操作性能等

3. B 型超声诊断仪

由于超声波频率高、波长短，可以像光那样沿直线传播，所以就有可能向某确定方向上发射，并顺利地人体组织里传播。当超声波遇到不同的介质交接面时会产生反射波。这些特点构成了超声仪器在医学领域广泛应用的基础。B 型超声诊断采用辉度调制（Brightness modulation）成像，简称 B 超。其基本原理是：在超声频率电脉冲激励下，用具有压电效应晶体的超声探头（probe）向人体发射一组超声波，按一定的方向进行扫描。探头接收回来的回声信号经过滤波、对数放大等信号处理。然后由数字扫描转换器（DSC，Digital Scanning Converter）进行数字变换形成数字信号，在 CPU 控制下进一步进行图像处理，再同图表形成电路和测量电路一起合成视频信号送给显示器，形成以不同辉度的光点强弱显示脏器及病变的二维切面黑白图像。根据监测其回声的延迟时间、强弱就可以判断脏器的距离及性质。



(a) 原理图



(b) 成像图

图 3.15 B 型超声诊断仪的工作原理与成像

B 超的主要性能指标包括：① 灰阶（灰度）：一般为 256 级。② 分辨率：分侧向分辨率、轴向分辨率，以 mm 为单位。③ 功能：有 A 型（将超声探头置于人体表面某一点上，声束射入体内，由组织界面返回的信号幅值显示于屏幕上，屏幕的横坐标表示超声波的传播时间，即探测深度，纵坐标则表示回波脉冲的幅度（Amplitude））、M 型（可获得界面运动（Motion）的轨迹图）、D 型（利用多普勒效应（Doppler effect），检测出人体内运动组织的信息）、多种测量能力（距离、面积、周长、体积

等)、多幅图像存储、多段灵敏度时间控制曲线 (Sensitivity-Time Control, STC) 自由控制、动态聚焦、可配宽频探头等。④ 探头: 可配多种探头, 如心脏、腹部、食道探头等。⑤ 图像处理: 黑白翻转、图像边缘处理、平滑处理、 γ 修正等。

目前出现的彩色超声波诊断仪 (简称“彩超”) 是指在 B 超的基础上, 根据多普勒效应以实时伪彩色编码 (将在第 10 章中介绍) 显示血流的一种方法, 不同的颜色代表血流的方向、速度和性质。而三维成像也是在 B 型法的基础上用专用探头对脏器进行容积式扫描, 然后利用计算机进行三维重建, 获得三维图像。四维超声是在三维超声基础上加上时间参数, 形成三维立体电影回放图像。

3.3.2 图像输出设备

数字图像的显示是图像数字化的逆过程——数/模转换 (D/A), 其目的在于将数字图像转化为适合于人眼观察的模拟形式。在多媒体技术中, 显示器和其他图像输出设备 (如打印机、胶片记录仪、静电绘图仪 (见图 3.16) 等) 都可以视为输出显示媒体。显示器是典型的暂时显示设备 (volatile display), 而打印机是永久显示设备 (permanent display)。图 3.16 是绘图仪的外形图。

1. 暂时显示设备

常见的暂时显示设备主要有采用光栅扫描的阴极射线管 (Cathode Ray Tube, CRT) 显示器和液晶显示器 (LCD)、等离子显示器、LED 显示器等。采用 CRT 显示是实现最早、应用最为广泛的一种显示技术。CRT 显示器的工作原理与普通电视机的显像管一样, 依靠反映图像信号强度的电子束轰击荧光屏产生在空间上亮度变化的光学图像。由于控制阴极的图像信号是随时间变化的一维信号, 因此必须通过扫描转化为按空间变化的图像, 所以显卡中的帧存储器的图像数据必须不断地刷新才能维持静态图像或视频图像的显示。CRT 显示器具有亮度高、色彩鲜艳且价格较低等优点, 但体积大、重量重、功耗大且存在电磁辐射。

液晶 (Liquid Crystal Display, LCD) 显示器是基于液晶电光效应的显示器件, 是很有发展前途的显示器。液晶是介于液体和固体之间的物质, 本身不发光, 工作性质受温度影响很大, 其工作温度为 $-55^{\circ}\text{C} \sim +77^{\circ}\text{C}$ 。液晶显示器利用液晶的光电效应, 即液晶分子的排列在电场作用下发生变化, 影响其液晶单元的透光率或反射率, 从而产生具有不同灰度层次及颜色的图像。液晶显示器具有体积小、重量轻、电压低、功耗低、不闪烁、无辐射等优点, 因而笔记本电脑、个人计算机都使用 LCD 显示器。LCD 显示器的工作原理基于液晶的物理特性: 在通电时液晶有秩序地排列, 光线容易通过; 不通电时, 排列则变得混乱, 阻止光线通过。显示器由段显示方式的字符段显示器件、矩阵显示方式的字符/图形/图像显示器件、矩阵显示方式的大屏幕液晶投影电视液晶屏等组成。LCD 显示器主要有扭曲向列型液晶显示器 (TN 型 LCD)、TFT 型液晶显示器 (TFT 型 LCD) 和高分子散布型液晶显示器 (PDLC 型 LCD) 三种, 其中 TFT 型 LCD 显示器已成为彩色液晶平板显示器的主流产品。

等离子显示器 (Plasma Display Panel, PDP) 是采用新一代等离子平面屏幕技术的显示设备。这种显示器的主要特点是提供大屏视角, 图像清晰逼真, 可满足高清晰度电视 (HDTV) 的要求; 亮度高; 在室外及普通居室光线下均可视; 屏幕轻薄, 厚度仅有厘米级, 便于安装。彩色 PDP 采用的数字灰度技术可使图像灰度超过 256 级, 可以显示 16 位或 24 位真彩色图像。

利用发光二极管 (Light-Emitting Diode, LED) 技术, 近年来 LED 背光技术迅速普及, 人们已制造出比 LCD 更薄、更亮、更清晰的显示器。LED 与 LCD 的功耗比大约为 1:10, 有机 LED 显示屏的单个元素反应速度是 LCD 液晶屏的 1000 倍。更高的刷新速率使得 LED 在图像和视频方面有更好的性能表现, 多幅显示屏还可以进行联网播出。



图 3.16 绘图仪

投影机（projector）主要通过三种显示技术实现：CRT 投影技术、LCD 投影技术以及 DLP 投影技术。

（1）CRT 投影技术

CRT 三枪投影机采用 CRT 作为成像器，在大屏幕上显示出彩色图像。光学系统与 CRT 管组成投影管。由于使用内光源，也叫主动式投影方式。CRT 技术成熟，显示的图像色彩丰富，还原性好，具有丰富的几何失真调整能力。但其重要技术指标图像分辨率与亮度相互制约，直接影响 CRT 投影机的亮度值，其亮度值一般在 300 lm 以下。另外，CRT 投影机操作复杂，特别是会聚调整烦琐，机身体积较大，适合安装于环境光较弱、相对固定的场所。

（2）LCD 投影技术

LCD 投影机分为液晶板和液晶光阀两种。液晶光阀投影机采用 CRT 管和液晶光阀作为成像器件，采用外光源，有效解决了图像分辨率与亮度间的矛盾，也叫被动式投影方式。一般的光阀主要由 3 部分组成：光电转换器、镜子、光调制器。通过 CRT 输出的光信号照射到光电转换器上，将光信号转换为持续变化的电信号；外光源产生一束强光，投射到光阀上，由内部的镜子反射，能过光调制器，改变其光学特性，紧随光阀的偏振滤光片，将滤去其他方向的光，而只允许与其光学缝隙方向一致的光通过，这个光与 CRT 信号相复合，投射到屏幕上。它是目前为止亮度、分辨率最高的投影机，亮度可达 6000 lm，分辨率为 2500×2000，适用于环境光较强、观众较多的场合，如超大规模的指挥中心、会议中心及大型娱乐场所，但其价格高，体积大，光阀不易维修。液晶板投影机的成像器件是液晶板，也是一种被动式的投影方式。利用外光源金属卤素灯或 UHP（冷光源），若是 3 块 LCD 板的设计，则把强光通过分光镜形成 RGB 三束光，分别透射过 RGB 三色液晶板；信号源经过模数转换，调制加到液晶板上，控制液晶单元的开启、闭合，从而控制光路的通过断，再经镜子合光，由光学镜头放大，显示在大屏幕上。目前，市场上常见的液晶投影机比较流行单片设计（LCD 单板，光线不用分离），这种投影机体积小，重量轻，操作、携带方便，价格低廉。但其光源寿命短，色彩不很均匀，分辨率较低，最高分辨率为 1024×768，多用于临时演示或小型会议。投影机实现了数字化调制信号，但由于液晶本身的物理特性，其响应速度慢，且性能随使用时间推移会有所下降。

（3）DLP 投影技术

数字光处理器（Digital Light Processor, DLP）也称数码投影机，该投影技术实现了数字信息的显示。它以数字微反射器（Digital Micromirror Device, DMD）作为光阀成像器件。根据所用 DMD 的片数，DLP 投影机可分为：单片机、两片机、三片机。在 DMD 块上，每一个像素的面积为 $16\mu\text{m}\times 16\mu\text{m}$ ，间隔为 $1\mu\text{m}$ 。DLP 投影机具有数字系统的优点，图像灰度等级达 256~1024 级，色彩达 $256^3\sim 1024^3$ 种，无图像噪声，对比度和清晰度高，分辨率已经达到 1280×1024，画面均匀，色彩锐利，画面质量稳定。其还具有反射优势。反射式 DMD 使成像器件的总光效率达 60% 以上，亮度可达 2000 lm 以上，随意变焦，调整十分便利。

2. 永久显示设备

目前常用的图像输出设备主要有各种打印机，如激光打印机、喷墨打印机等，它们已经广泛在办公自动化领域获得应用。由于图像处理的要求，要求打印机的输出像素的密度在 300 dpi 以上，通常为 600 dpi。彩色图像处理的结果还要用彩色打印机输出。

对于大幅面、高质量等具有专业要求的图像输出，则需要其他的图像显示设备，如胶片记录仪、静电绘图仪等。胶片记录仪将图像输出到感光胶片上，一般胶片记录仪可以输出 2 K（即分辨率为 2 000 线）的图像。35 mm 胶片的分辨率至少达到 4 K。静电绘图仪是一种光栅扫描设备，分辨率可达 800 dpi。彩色静电绘图仪采用高质量的黄、紫、青、黑四色墨水和纸张，可在不同的位置形成 4 000 多种色彩的图案，可产生比彩色照片还好的效果。

3.4 实验：图像的数字化

1. 实验目的

- (1) 了解静止图像的数字化原理。
- (2) 掌握常用图像数字化设备的使用：摄像机、扫描仪、数码相机等。

2. 实验主要仪器设备

- (1) 摄像机。
- (2) 台式机或笔记本电脑。
- (3) Pinnacle System（含视频采集与编辑软件的系统，也可用性能类似的采集装置代替）。
- (4) 单反数码相机或普通数码相机。
- (5) 扫描仪。

3. 实验原理

自然界的图像是以模拟形式出现的。只有将模拟图像数字化后，才能便于计算机灵活地对图像进行各种处理。典型的微机图像处理系统如图 3.17 所示。图像采集板的基本功能是对摄像机输出的模拟图像进行采样和量化而输出数字图像。

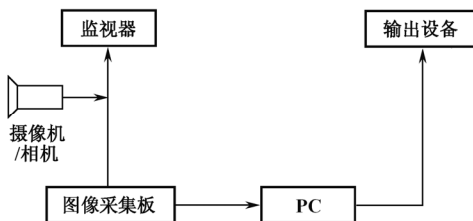


图 3.17 微机图像处理系统

4. 实验内容和步骤

(1) 实验内容

- ① 构建微机图像系统。
- ② 采集视频图像和静态图像。
- ③ 编辑静态图像。

(2) 实验步骤

- ① 将摄像机、视频采集卡、计算机相连，构成硬件系统。
- ② 安装 Pinnacle System 的软件——Pinnacle Studio 8。
- ③ 用 Pinnacle System 采集视频。

④ 用 Pinnacle Studio 8 的“帧抓取器”获取静态图像。一旦被抓取的帧保存到磁盘上，就可以在其他软件程序中使用或在图像编辑软件中编辑。

- ⑤ 用数码相机摄取感兴趣的景物。
- ⑥ 用扫描仪获取几种典型画面：二值文本图像、灰度图像、彩色图像。
- ⑦ 编程实现图 3.5 中不同采样率图像的显示效果。
- ⑧ 编程实现图 3.6 中不同量化等级图像的显示效果。

5. 实验报告要求

- (1) 说明图像采集系统的组成和原理。
- (2) 记录摄取的静态图像。
- (3) 记录编辑静态图像的有关操作。
- (4) 回答思考题。
- (5) 心得和体会。

6. 预习要求

- (1) 了解摄像机、数码相机、扫描仪的基本原理。
- (2) 了解图像采集板的框图原理。

7. 思考题

- (1) 为什么同一图像采用不同格式时的数据量不同？
- (2) 叙述视频图像与静态图像之间的关系。

本章小结

本章的数学基础是数字图像的描述方法，重点是二维采样和量化的概念，难点是二维采样定理和向量量化方法。读者在学习本章时，应该将输入设备的采样量化过程与输出设备的恢复过程联系起来，对数字图像处理系统的输入/输出设备的组成原理和性能指标也应该有深刻的了解。

思考题与习题

3.1 光圈越大, 摄取的画面是否越亮? 光圈的数字越小是否代表光圈越小? 2.8 的光圈与 5.6 的光圈, 哪个画面更亮? 光圈越小, 画面是否越有立体感?

3.2 换一幅图像, 重新演算图 3.4 中实际图像量化的运行结果。

3.3 如何合理选取采样率和量化等级数?

3.4 讨论基于二维采样定理, 根据图像采集设备的分辨率计算最大像素数。

3.5 验证图 3.8 一维向量量化的误差 $d(X, C_i)$ 的正确性。

3.6 计算图 3.9 图像 X 与图 3.10 码书 C 码字的误差, 验证选择码字 Y_1 作为 X 的二维向量量化编码的合理性。

3.7 根据“电视行 (TV line)”或扫描线的概念, 摄像机的清晰度常用扫描线的线数表示。如果画面是 4:3 的, 540 线的像素分辨率多少? 720p 是 720 线 (p 代表逐行扫描, i 是隔行扫描), 帧型为 16:9, 像素分辨率为 1280×720。

3.8 摄像机和数码相机分辨率受到哪些因素限制? 是否越高越好? 如何进一步提高分辨率?

3.9 扫描仪的光学分辨率是 600×1200 线, 一个具有 5000 个感光单元的 CCD 器件, 用于 A4 幅面扫描仪, A4 幅面的纸张宽度是 21 cm, 该扫描仪的光学分辨率是多少 dpi?

第4章 图像变换与二维数字滤波

内容提要

正交变换和二维数字滤波广泛应用于图像增强、图像复原、特征提取、图像编码与压缩及形状分析等方面。本章主要介绍图像处理中常用的二维离散变换的定义、性质、实现方法及应用,包括离散傅里叶变换(DFT)、离散余弦变换(DCT)、离散沃尔什-哈达玛变换(DWT)、K-L变换(KLT)、离散小波变换(DWT)及其应用。在图像变换的基础上,讨论二维数字滤波的定义、二维数字滤波器的设计及实现方法。

知识要点

- 余弦型变换: 傅里叶变换和余弦变换。
- 方波型变换: 沃尔什-哈达玛变换。
- 基于特征向量的变换: K-L变换。
- 从哈尔变换、短时傅里叶变换到小波变换。
- 各种变换的定义和有关快速算法及实现方法。
- 二维数字滤波的定义、二维数字滤波器的设计及实现方法。

教学建议

- 建议安排6学时。
- 本章数学公式较多,变换和滤波的结果不像原始图像那样可以直接被人眼所感知,所以讲授和学习各种变换和滤波时应当注意它们的物理意义。在后续章节的学习中要注意各种变换和滤波的实际应用。
- 本章的先修知识包括一维积分变换、傅里叶变换、随机过程、小波变换等。

4.1 二维离散傅里叶变换（DFT）

图像处理的计算方法一般为线性的，处理后的输出图像可以看成是输入图像阵列各元素的加权组合，这比非线性方法要简单得多。即便如此，由于实际的图像矩阵很大，其计算复杂度可想而知。况且图像的一些重要特征，如频谱等在空间域（spatial domain）难以直接观测，频域为我们观察图像提供了新的视角。因此，各种线性变换方法可对图像获得更有效的处理。

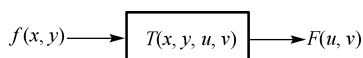


图 4.1 二维正交变换

图像变换一般为线性正交变换，如图 4.1 所示， $f(x, y)$ 、 $F(u, v)$ 分别是原始图像和变换后的图像。 $T(x, y; u, v)$ 为变换基函数（transform basis function）。不同的变换基形成不同的正交变换。

傅里叶变换是最基本的正交变换。二维离散傅里叶变换是在一维离散傅里叶变换基础上扩展而来的，在理论和实用上具有重要的价值。由于在数字信号处理课程中已详细介绍了一维傅里叶变换，本节直接引入二维傅里叶变换。图像经过变换后，成为变换域的矩阵，空间域图像的处理问题转化为变换域的滤波（transform domain filtering）处理，这时处理速度更快，并可利用二维滤波技术对图像进行灵活的处理。

4.1.1 二维连续傅里叶变换

二维连续函数 $f(x, y)$ 的傅里叶变换定义如下。

设 $f(x, y)$ 是两个独立变量 x 和 y 的函数，且在 $\pm\infty$ 上绝对可积，即积分

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)| dx dy < \infty \quad (4.1)$$

则定义积分

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (4.2)$$

为二维连续函数 $f(x, y)$ 的傅里叶变换，并定义

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv \quad (4.3)$$

为 $F(u, v)$ 的逆变换。我们把 $f(x, y)$ 和 $F(u, v)$ 称为傅里叶变换对。

【例 4.1】 求图 4.2 所示函数 $f(x, y) = \begin{cases} A, & |x| \leq X, |y| \leq Y \\ 0, & \text{其他} \end{cases}$ 的傅里叶变换。

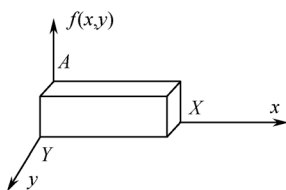
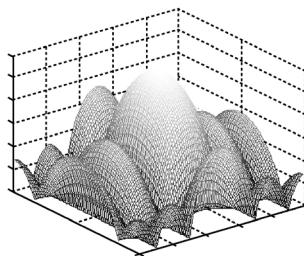
【解】 将函数（其图形如图 4.1）代入式（4.2）中，得

$$\begin{aligned} F(u, v) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy = A \int_0^X e^{-j2\pi ux} dx \int_0^Y e^{-j2\pi vy} dy \\ &= AXY \cdot \frac{\sin(\pi uX)}{\pi uX} e^{-j\pi ux} \cdot \frac{\sin(\pi vY)}{\pi vY} e^{-j\pi vy} \end{aligned}$$

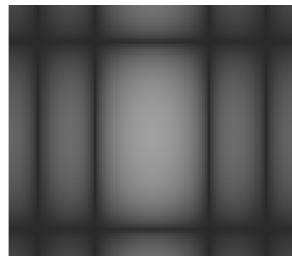
其幅度谱为

$$|F(u, v)| = AXY \left| \frac{\sin(\pi uX)}{\pi uX} \right| \left| \frac{\sin(\pi vY)}{\pi vY} \right|$$

相应的图形如图 4.3（a）所示，图 4.3（b）是用灰度表示的幅度谱。

图 4.2 二维信号 $f(x, y)$ 

(b) 频谱图的三维表示



(b) 频谱的灰度图

图 4.3 信号的频谱图

4.1.2 二维离散傅里叶变换

连续傅里叶变换无法用数字计算机实现，而离散傅里叶变换建立了离散时间域与离散频率域之间的关系，物理意义强且有快速算法——FFT（快速傅里叶变换），因而在信号分析与图像处理领域具有很大的实用价值。另外，二维傅里叶变换是在一维傅里叶变换基础上建立的，其定义、性质和方法只要在一维变换基础上适当地推广即可完成。

尺寸为 $M \times N$ 的离散函数 $f(x, y)$ 的二维离散傅里叶变换（Discrete Fourier Transform, DFT）如下

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (4.4)$$

$f(x, y)$ 可以通过对 $F(u, v)$ 求傅里叶逆变换获得，其表达式如下

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)} \quad (4.5)$$

式中， $x=0, 1, 2, \dots, M-1$ ； $y=0, 1, 2, \dots, N-1$ 。式（4.4）和式（4.5）构成了二维离散傅里叶变换对。变量 u 和 v 是频率变量， x 和 y 是空间变量。

$F(u, v)$ 即为离散信号 $f(x, y)$ 的频谱，通常是复数

$$F(u, v) = R(u, v) + jI(u, v) \quad (4.6)$$

式中， $R(u, v)$ 和 $I(u, v)$ 分别是 $F(u, v)$ 的实部和虚部。它的模和相角分别为

$$|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{1/2} \quad (4.7)$$

$$\phi(u, v) = \arctan \left[\frac{I(u, v)}{R(u, v)} \right] \quad (4.8)$$

称 $|F(u, v)|$ 和 $\phi(u, v)$ 为 $f(x, y)$ 的幅度谱和相位谱。

在实际工程应用中分析幅度谱较多，习惯上也常把幅度谱称为频谱。使用 DFT 进行图像处理时，有如下特点：

① 频谱的直流成分为 $F(0, 0) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)$ ，说明在频谱原点的傅里叶变换 $F(0, 0)$ 等于图像的平均灰度级。

② 幅度谱 $|F(u, v)|$ 关于原点对称，即 $F(u, v) = F(-u, -v)$ 。

③ 图像 $f(x, y)$ 平移后，幅度谱不发生变化，仅有相位发生变化。

4.1.3 二维离散傅里叶变换的性质

二维离散傅里叶变换是一种线性算子，其中不少性质与一维傅里叶变换的性质很相似。表 4.1 给出了一些基本性质。

表 4.1 二维傅里叶变换的基本性质

性 质	表 达 式
1. 线性	$\text{DFT}[a_1 f_1(x, y) + a_2 f_2(x, y)] = a_1 \text{DFT}[f_1(x, y)] + a_2 \text{DFT}[f_2(x, y)]$
2. 变换可分离性	$F(u, v) = F_x\{F_y[f(x, y)]\} = F_y\{F_x[f(x, y)]\}$ $f(x, y) = F_u^{-1}\{F_v^{-1}[F(u, v)]\} = F_v^{-1}\{F_u^{-1}[F(u, v)]\}$
3. 比例性质	$f(ax, by) \leftrightarrow \frac{1}{ ab } F\left(\frac{u}{a}, \frac{v}{b}\right)$
4. 空间位移	$f(x - x_0, y - y_0) \leftrightarrow F(u, v) e^{-j2\pi(u x_0 + v y_0)/N}$
5. 频移（调制）	$f(x, y) e^{j2\pi(u_0 x + v_0 y)/N} \leftrightarrow F(u - u_0, v - v_0)$
6. 对称性	若 $f(x, y) = f(-x, -y)$ ，则 $F(u, v) = F(-u, -v)$
7. 共轭对称性	$f^*(x, y) \leftrightarrow F^*(-u, -v)$
8. 频率域卷积	$f_c(x, y) g_c(x, y) \leftrightarrow F_c(u, v) * G_c(u, v)$ ，其中 f_c 和 g_c 是 f 和 g 的周期延拓
9. 空间域卷积	$f_c(x, y) * g_c(x, y) \leftrightarrow F_c(u, v) G_c(u, v)$ ，其中 F_c 和 G_c 是 F 和 G 的周期延拓
10. 平均值	$F(0, 0) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)$
11. 180° 旋转	$\text{DFT}\{\text{DFT}[f(x, y)]\} = f(-x, -y)$
12. 周期性	$f(x, y) = f(x + mN, y + nN)$, $F(u, v) = F(u + mN, v + nN)$ ，式中 $m, n = 0, \pm 1, \pm 2, \dots$

下面对其中的某些性质做详细的介绍。

1. 变换可分离性

可分离性是指二维 DFT 可以用两个可分离的一维 DFT 之乘积形式表示

$$F(u, v) = \frac{1}{M} \sum_{x=0}^{M-1} e^{-j2\pi ux/M} \frac{1}{N} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi vy/N} = \frac{1}{M} \sum_{x=0}^{M-1} F(x, v) e^{-j2\pi ux/M} \quad (4.9a)$$

式中，

$$F(x, v) = \frac{1}{N} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi vy/N} \quad (4.9b)$$

对于每个 x 值，当 $v = 0, 1, 2, \dots, N-1$ 时， $F(x, v)$ 是沿着 $f(x, y)$ 的一行所进行的一维傅里叶变换。然后再将 $F(x, v)$ 逐列进行变换，就得到了 $F(u, v)$ 。这个步骤如图 4.4 所示。这样我们可以得到结论：二维变换可以通过先进行行变换再进行列变换的两次一维变换来实现。同样，也可以通过先求列变换再求行变换得到二维傅里叶变换。

2. 周期性、共轭对称性及频谱中心化

设 a 和 b 为整数，离散傅里叶变换有如下的周期性

$$F(u, v) = F(u + aM, v) = F(u, v + bN) = F(u + aM, v + bN) \quad (4.10a)$$

同样，逆变换也具有周期性

$$f(x, y) = f(x + aM, y) = f(x, y + bN) = f(x + aM, y + bN) \quad (4.10b)$$

如图 4.5 所示, $f(x, y)$ 的 DFT 频谱是在水平方向以 N 为周期、在垂直方向以 M 为周期的周期图形。

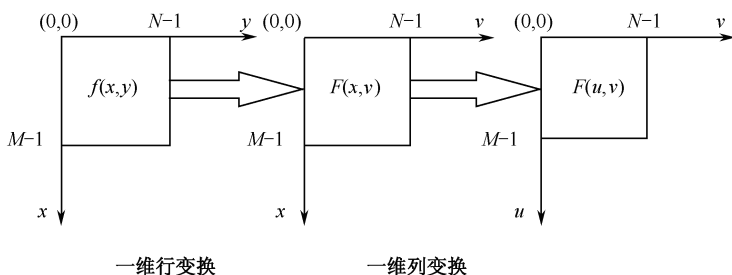


图 4.4 用两次一维 DFT 计算二维 DFT

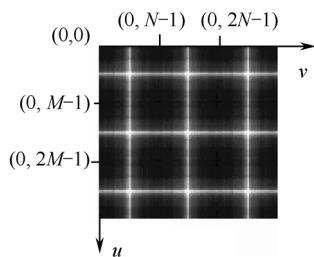


图 4.5 典型图像的频谱图

由于图像函数 $f(x, y)$ 一般为实函数，DFT 的共轭对称性可以表示为

$$F(u, v) = F^*(-u, -v) \quad (4.11a)$$

和

$$|F(u, v)| = |F^*(-u, -v)| \quad (4.11b)$$

在实际应用中，周期性和共轭对称性给我们带来了许多方便。首先来看一维的情况。

设一个矩形函数

$$f(x) = \begin{cases} A, & 0 \leq x \leq X \\ 0, & \text{其他} \end{cases}$$

可以求出它的傅里叶变换为

$$F(u) = AX \frac{\sin \pi u X}{\pi u X} e^{-j\pi u X}$$

进而可得幅度谱为

$$|F(u)| = AX \left| \frac{\sin \pi u X}{\pi u X} \right|$$

其幅度谱图形如图 4.6 (a) 所示。 $|F(u)|$ 具有长度为 N 的周期，由共轭对称性 $|F(u)| = |F(-u)|$ 可知频谱关于原点中心对称，因此，在 $-N/2$ 到 $N/2$ 之间可以看到一个完整的频谱，而离散傅里叶变换的求和区间是 $[0, N-1]$ ，在这个区间内，频谱是由两个背靠背的半周期组成的。要显示一个完整的周期，必须将变换的原点移至点 $u = N/2$ ，如图 4.6 (b) 所示。根据定义，我们有

$$F(u + N/2) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j\frac{2\pi}{N}x(u+N/2)} = \frac{1}{N} \sum_{x=0}^{N-1} (-1)^x f(x) e^{-j\frac{2\pi}{N}xu} \quad (4.12)$$

由此可以看到，在进行傅里叶变换之前用 $(-1)^x$ 乘以输入信号 $f(x)$ ，便可以在一个周期的变换中 ($u = 0, 1, 2, \dots, N-1$) 求得一个完整的频谱。

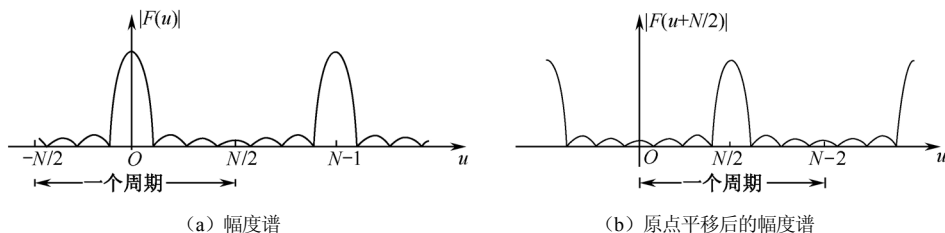


图 4.6 频谱图

同样的结论可以推广到二维情况。在进行傅里叶变换之前，用 $(-1)^{x+y}$ 乘以输入的图像函数，则有

$$\text{DFT}[f(x,y)(-1)^{x+y}] = F(u-M/2, v-N/2) \quad (4.13)$$

式 (4.13) 说明 $f(x,y)(-1)^{x+y}$ 的傅里叶变换的原点被设置在 $u = M/2$ 和 $v = N/2$ 上。($M/2, N/2$) 是二维 DFT 设置的 $M \times N$ 区域的中心。我们将此频率域的范围指定为频率矩形，它从 $u = 0$ 到 $u = M-1$ ，从 $v = 0$ 到 $v = N-1$ 。为了确保移动后的坐标为整数，要求 M 和 N 为偶数。

图 4.7 (a) 是一个在黑色背景上叠加一个白色矩形的图像。图 4.7 (b) 是中心化前图 4.7 (a) 的频谱图，变换中心在左上角，而 4 个二维周期分量从图像的 4 个角背靠背地扩展到中心，高频成分位于中心部位。图 4.7 (c) 是中心化后的频谱图，此时变换的零频率成分在中心，高频成分位于周围。可见中心化有助于我们观察频谱。

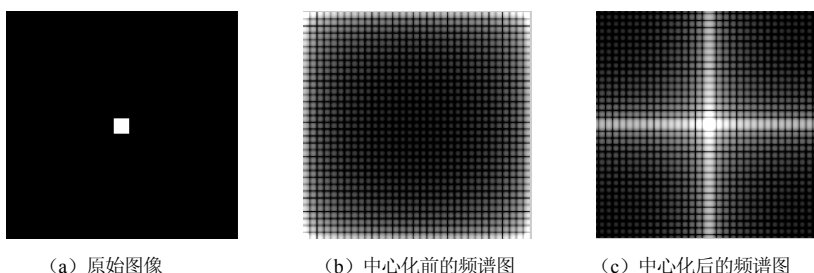


图 4.7 图像频谱的中心化

3. 离散卷积定理

卷积定理和相关定理研究的都是两个函数的 DFT 之间的关系，也构成了空间域和频率域之间的基本关系。

设 $f(x,y)$ 和 $g(x,y)$ 是大小分别为 $A \times B$ 和 $C \times D$ 的两个数组，卷积定理要求求两个函数的傅里叶变换，再求两个傅里叶变换乘积的逆变换来求这两个函数的卷积。而这样解得的卷积是周期函数，因而求出的卷积是循环卷积。为了防止卷积后产生交叠误差 (wraparound error)，使得循环卷积和线性卷积的结果保持一致，可以先对两个函数补零。回顾数字信号处理课程中一维循环卷积的知识，对于大小分别为 A 和 C 的两个数组 $f(x)$ 和 $g(x)$ ，它们的循环卷积 $f(x) * g(x)$ 的长度为 $N = A + C - 1$ ，因此，可以把两个被卷积函数的长度均扩展到 N ，并在原函数定义区间以外的部分补零。于是，在一个周期为 N 的卷积计算中，便不会发生边缘效应，从而求得的循环卷积结果就等于所要求的线性卷积。

同样，在利用离散卷积定理计算二维离散卷积时，也必须对被卷积函数进行延拓和补零。如果被卷积函数 $f(x,y)$ 和 $g(x,y)$ 的大小分别为 $A \times B$ 和 $C \times D$ ，则延拓后的函数应为

$$f_e(x,y) = \begin{cases} f(x,y), & 0 \leq x \leq A-1, 0 \leq y \leq B-1 \\ 0, & A \leq x \leq M-1, B \leq y \leq N-1 \end{cases} \quad (4.14)$$

$$g_e(x,y) = \begin{cases} g(x,y), & 0 \leq x \leq C-1, 0 \leq y \leq D-1 \\ 0, & C \leq x \leq M-1, D \leq y \leq N-1 \end{cases} \quad (4.15)$$

式中， $M \geq A + C - 1$ ， $N \geq B + D - 1$ 。于是所求的离散卷积为

$$f_e(x,y) * g_e(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m,n) g_e(x-m, y-n) \quad (4.16)$$

式中， $x = 0, 1, 2, \dots, M-1$ ， $y = 0, 1, 2, \dots, N-1$ ； $M = A + C - 1$ ， $N = B + D - 1$ 。

二维离散卷积定理为

$$\text{DFT}[f_c(x,y) * g_c(x,y)] = F_c(u,v) \cdot G_c(u,v) \quad (4.17)$$

式中, $F_c(u,v)$ 和 $G_c(u,v)$ 分别是 $f_c(x,y)$ 和 $g_c(x,y)$ 的傅里叶变换。

DFT 是一种基本和重要的正交变换。为了提高计算效率, 应用时往往采用二维 FFT 实现。而一般的正交变换图像经过对数变换后更便于观察。MATLAB 采用 `fft2` 和 `ifft2` 分别进行二维 DFT 变换和二维 DFT 逆变换, 采用 `fftshift` 将直流分量移到频谱图的中心以便于观察。

【例 4.2】 用 MATLAB 实现图像的傅里叶变换。为了增强显示效果, 程序中用对数对频谱的幅度进行压缩, 然后将频谱幅度的对数值用在 0~10 之间的值进行显示。

【解】 MATLAB 程序如下:

```
I = imread('pout.tif');           %读入图像
imshow(I);                         %显示图像
F1 = fft2(I);                     %计算二维傅里叶变换
figure, imshow(log(abs(F1)+1), [0 10]); %显示对数变换后的频谱图
F2 = fftshift(F1);                %将直流分量移到频谱图的中心
figure, imshow(log(abs(F2)+1), [0 10]); %显示对数变换后中心化的频谱图
```

程序运行的结果如图 4.8 所示。

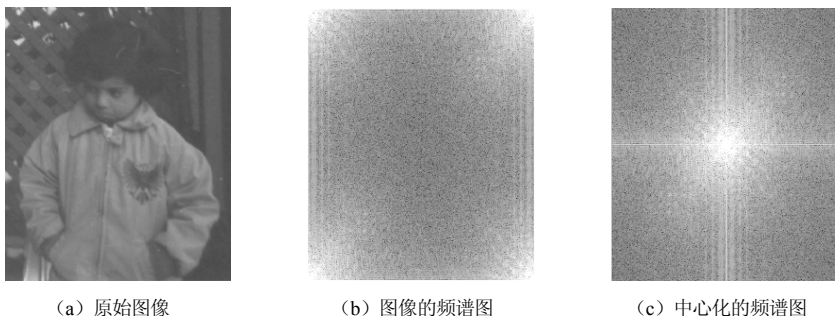


图 4.8 傅里叶变换

4.2 二维离散余弦变换 (DCT)

离散余弦变换 (Discrete Cosine Transform, DCT) 类似于 DFT, 但为实数变换, 是简化傅里叶变换的重要方法, 在图像数据压缩中具有重要应用。将信号延拓为实偶函数, DCT 相当于一个长度是信号长度两倍的 DFT, 而一个实偶函数的傅里叶变换仍然是一个实偶函数。同样, 如果考虑 DFT 的虚数部分, 则对应于离散正弦变换 (DST)。

4.2.1 一维离散余弦变换

一维傅里叶级数展开式中, 若被展开的函数是实偶函数, 那么其傅里叶级数只有余弦函数。由此可以将一个信号通过对折延拓成实偶函数, 然后进行傅里叶变换, 我们就可用 $2N$ 点的 DFT 来产生 N 点的 DCT。一维 DCT 的推导过程在一般数字信号处理教程中均有介绍, 这里直接给出正、逆变换的公式。

定义 $f_c(n)$ 的离散余弦变换为

$$F(k) = C(k) \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} f(n) \cos \left[\frac{\pi(2n+1)k}{2N} \right] \quad (4.18)$$

式中, 引入常量 $C(k)$ 是为了保证变换基的规范正交性, 且

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k=0 \\ 1, & 1 \leq k \leq N-1 \end{cases} \quad (4.19)$$

同样, DCT 逆变换为

$$f(n) = \frac{1}{\sqrt{N}} C(0) + \sqrt{\frac{2}{N}} \sum_{u=1}^{N-1} F(u) \cos \left[\frac{\pi(2n+1)u}{2N} \right] \quad (4.20)$$

4.2.2 二维离散余弦变换

对于二维信号, 同样可以推出它的离散余弦变换为

$$F(u, v) = C(u)C(v) \sqrt{\frac{2}{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{\pi}{M} u \left(x + \frac{1}{2} \right) \right] \cos \left[\frac{\pi}{N} v \left(y + \frac{1}{2} \right) \right] \quad (4.21)$$

式中 $C(u)$ 、 $C(v)$ 的定义同 $C(k)$ 。

DCT 逆变换为

$$f(x, y) = \sqrt{\frac{2}{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} C(u)C(v) F(u, v) \cos \left[\frac{\pi}{M} u \left(x + \frac{1}{2} \right) \right] \cos \left[\frac{\pi}{N} v \left(y + \frac{1}{2} \right) \right] \quad (4.22)$$

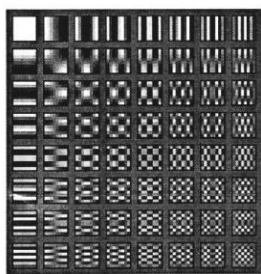


图 3.9 8×8 的基本图像

DCT 的逆变换式可以理解为: 任意大小为 $M \times N$ 的图像 $f(x, y)$ 可以写成 $M \times N$ 个如式 (4.23) 的函数的线性组合形式, 即

$$C(u)C(v)F(u, v) \cos \left[\frac{\pi}{M} u \left(x + \frac{1}{2} \right) \right] \cos \left[\frac{\pi}{N} v \left(y + \frac{1}{2} \right) \right] \quad (4.23)$$

这些函数被称为 DCT 的基本函数 (图像)。DCT 系数 $F(u, v)$ 可以视为对应函数的权值。例如 8×8 的图像 $f(x, y)$ 可以视为 64 个基本图像的线性组合, 其基本图像如图 4.9 所示。其中, $(0, 0)$ 位置的元素就是直流分量, 矩阵中的其他元素根据其位置表示不同频率的交流分量。

MATLAB 采用 `dct2` 和 `idct2` 分别进行二维 DCT 变换和二维 DCT 逆变换。

4.2.3 二维 DCT 的应用

二维 DCT 常用于二维信号处理, 典型应用是对静止图像和运动图像进行性能优良的有损数据压缩。在静止图像编码标准 JPEG、运动图像编码标准 MJPEG 和 MPEG 等标准中都使用了 8×8 块的离散余弦变换, 并将结果进行量化之后进行熵编码。DCT 具有很强的能量集中在频谱的低频部分的特性, 而且当信号具有接近马尔可夫过程 (Markov process) 的统计特性时, DCT 的去相关性接近于具有最优去相关性的 K-L 变换 (Karhunen-Loève 变换) 的性能。

另外, 改进的离散余弦变换 (Modified Discrete Cosine Transform, MDCT) 对交叠的数据进行 DCT, 有助于避免由于区块边界所产生的多余数据, 被用在高级音频编码 (Advanced Audio Coding, AAC)、Ogg Vorbis、AC-3 和 MP3 音频压缩中。

【例 4.3】应用 MATLAB 实现图像的 DCT 变换。

【解】MATLAB 程序如下:

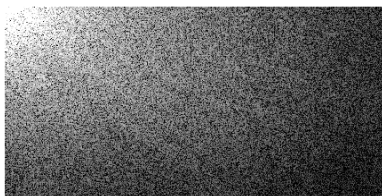
```
I = imread('wpeppers2.png');
J = rgb2gray(I); %转换彩色图像为灰度图像
```

```
subplot(1,2,1),imshow(J);           %显示原灰度图像
K = dct2(J);                         %对图像做 DCT 变换
subplot(1,2,2),imshow(log(abs(K))+1,[0 10]); %显示 DCT 变换结果
```

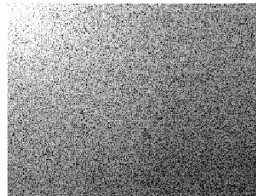
程序运行结果如图 4.10 所示。



(a) wpeppers2 图像



(b) wpeppers2 图像的 DCT 系数



(c) 显微图像的 DCT 系数

图 4.10 离散余弦变换

对 SBS 改性沥青材料的显微图像 (见图 2.20) 进行上述变换, 使图像的空间灰度分布转换成在频域图像灰度变化的程度。对图像来说, 图像边缘是灰度跳变部分, 变化较快, 在频域上对应高频分量; 反之, 图像变化平缓的部分则为低频分量。

4.3 二维离散沃尔什-哈达玛变换 (DHT)

前面介绍的像 DFT 和 DCT 一类的正交变换的基底函数, 基于复指数函数或余 (正) 弦函数, 因而占用较多的计算时间。在如通信等实时处理领域, 往往需要更为便利和有效的变换方法。选用方波信号或其变形作为变换基, 便于硬件实现且抗干扰性能好。通常, 这类变换中的许多乘法操作非常简单, 因而这类变换的计算速度很快。沃尔什-哈达玛 (Walsh-Hadamard) 变换就是其中的一种。

4.3.1 沃尔什变换

沃尔什函数系是函数值仅取+1 和-1 两值的非正弦型的标准正交完备函数系。由于二值正交函数与数字逻辑中的两个状态相对应, 所以非常便于计算机和数字信号处理器运算。图 4.11 显示了沃尔什函数系的前 10 个函数。

沃尔什函数有 3 种排列或编号方式, 即列率排列、佩利 (Paley) 排列 (自然排列) 和哈达玛 (Hadamard) 排列。这 3 种排列各有特点。沃尔什变换的排列方式为列率排列。与正弦波频率相对应, 非正弦波形可用列率描述。列率表示某种函数在单位区间上函数值为零的零点个数之半。

设 $N = 2^n$, 一维沃尔什变换核为

$$g(x, u) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)} \quad (4.24)$$

式中, $u, x = 0, 1, \dots, N-1$ 。 $b_k(z)$ 代表 z 的二进制表示的第 k 位值。沃尔什变换核是一个对称阵列, 其行和列是正交的。

一维沃尔什变换为

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)} \quad (4.25)$$

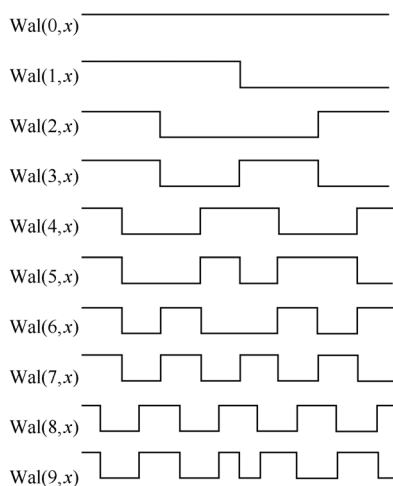


图 4.11 沃尔什函数系的前 10 个函数

逆变换为

$$f(x) = \sum_{u=0}^{N-1} W(u) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-i-1}(u)} \quad (4.26)$$

可见，正变换核、逆变换核仅相差一个常数项 $1/N$ 。计算正变换的任何算法同样适用于逆变换。二维沃尔什变换要求图像的大小为 $N = 2^n$ 。其正变换核为

$$g(x, y; u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i-1}(u) + b_i(y)b_{n-i-1}(v)]} \quad (4.27)$$

逆变换核为

$$h(x, y; u, v) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i-1}(u) + b_i(y)b_{n-i-1}(v)]} \quad (4.28)$$

二维沃尔什正变换和逆变换分别为

$$W(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i-1}(u) + b_i(y)b_{n-i-1}(v)]} \quad (4.29)$$

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} W(u, v) \prod_{i=0}^{n-1} (-1)^{[b_i(x)b_{n-i-1}(u) + b_i(y)b_{n-i-1}(v)]} \quad (4.30)$$

不难看出，沃尔什变换是可分离的和对称的。因此，二维沃尔什变换可以像二维 DFT 一样，分为两次一维沃尔什变换实现。

二维沃尔什变换的矩阵形式为

$$\mathbf{W} = \frac{1}{N^2} \mathbf{G} \mathbf{f} \mathbf{G} \quad (4.31)$$

式中， \mathbf{G} 为 N 阶沃尔什变换的核矩阵。二维沃尔什逆变换的矩阵形式为

$$\mathbf{f} = \frac{1}{N^2} \mathbf{G} \mathbf{W} \mathbf{G} \quad (4.32)$$

【例 4.4】 已知 $N=4$ ，求 $f(x)$ 的沃尔什变换。

【解】 由式 (4.28) 可得

$$\begin{aligned} W(0) &= \frac{1}{4} \sum_{x=0}^3 f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{2-i-1}(0)} = \frac{1}{4} [f(0) + f(1) + f(2) + f(3)] \\ W(1) &= \frac{1}{4} \sum_{x=0}^3 f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{2-i-1}(1)} = \frac{1}{4} [f(0) + f(1) - f(2) - f(3)] \\ W(2) &= \frac{1}{4} \sum_{x=0}^3 f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{2-i-1}(2)} = \frac{1}{4} [f(0) - f(1) + f(2) - f(3)] \\ W(3) &= \frac{1}{4} \sum_{x=0}^3 f(x) \prod_{i=0}^1 (-1)^{b_i(x)b_{2-i-1}(3)} = \frac{1}{4} [f(0) - f(1) - f(2) + f(3)] \end{aligned}$$

【例 4.5】 已知二维数字图像矩阵为 $\mathbf{f} = \begin{bmatrix} 2 & 5 & 5 & 2 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 2 & 5 & 5 & 1 \end{bmatrix}$ ，求此图像的二维 DWT，并反求 \mathbf{f} 。

【解】根据 $G = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$, 而 $W = \frac{1}{N^2} G f G$, 则可采用 MATLAB 程序求解 W 。

```
f = [2 5 5 2; 3 3 3 3; 3 3 3 3; 2 5 5 1];
G = [1 1 1 1; 1 1 -1 -1; 1 -1 -1 1; 1 -1 1 -1];
W = (1/16)*G*f*G
```

运行结果为:

```
W =
    3.1875    0.0625   -0.8125    0.0625
    0.0625   -0.0625    0.0625   -0.0625
    0.1875    0.0625   -0.8125    0.0625
    0.0625   -0.0625    0.0625   -0.0625
```

反求 f 的程序如下:

```
W = [3.1875    0.0625   -0.8125    0.0625;
     0.0625   -0.0625    0.0625   -0.0625;
     0.1875    0.0625   -0.8125    0.0625;
     0.0625   -0.0625    0.0625   -0.0625]
G = [1 1 1 1; 1 1 -1 -1; 1 -1 -1 1; 1 -1 1 -1];
f = G*W*G
```

运行结果为:

```
f =
     2     5     5     2
     3     3     3     3
     3     3     3     3
     2     5     5     1
```

从上例可见, W 左上角的数值相对高于右下角的数值, 所以二维 DHT 具有集中能量的特性。不难说明图像的数值越是均匀, 则变换后的数据越是集中在矩阵的左上角附近。因此, 利用 DHT 可以压缩图像信息。

4.3.2 哈达玛变换

如上所述, 沃尔什函数的三种排列方式各有特点。哈达玛排列定义简单, 存在从低阶到高阶的递推关系, 高阶矩阵可以由两个低阶矩阵之积求得, 便于快速计算, 实用性更好。采用哈达玛排列的沃尔什函数进行的变换称为沃尔什-哈达玛变换 (WHT) 或哈达玛变换。

哈达玛并变换矩阵是元素仅由+1 和-1 组成的正交方阵, 它的任意两行或两列都彼此正交, 即它们的对应元素之和为零。哈达玛变换核矩阵与沃尔什变换的差异仅仅是行的次序不同。

一维哈达玛变换核为

$$g(x, u) = \frac{1}{N} (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)} \quad (4.33)$$

式中, $u, x = 0, 1, \dots, N-1$ 。

一维哈达玛正变换为

$$H(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)} \quad (4.34)$$

式中, $u = 0, 1, \dots, N-1$ 。

一维哈达玛逆变换为

$$f(x) = \sum_{u=0}^{N-1} H(u) (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)} \quad (4.35)$$

式中, $x = 0, 1, \dots, N-1$ 。

二维哈达玛正变换和逆变换分别为

$$H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) (-1)^{\sum_{i=0}^{n-1} [b_i(x) b_i(u) + b_i(y) b_i(v)]} \quad (4.36)$$

和

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} H(u, v) (-1)^{\sum_{i=0}^{n-1} [b_i(x) b_i(u) + b_i(y) b_i(v)]} \quad (4.37)$$

可见, 二维哈达玛正、逆变换也具有相同的形式。哈达玛变换核是可分离的和对称的。因此, 二维的哈达玛正变换和逆变换都可通过两个一维变换实现。

最低阶的哈达玛矩阵为

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (4.38)$$

高阶哈达玛矩阵可以通过如下递推公式求得

$$\mathbf{H}_N = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & -\mathbf{H}_{N/2} \end{bmatrix} \quad (4.39)$$

例如, $N=4$ 的哈达玛矩阵为

$$\mathbf{H}_4 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \\ \mathbf{H}_2 & -\mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (4.40)$$

可见, 对于任意 N 阶哈达玛矩阵, 其元素仍然只含有 ± 1 , 而且可以根据上一阶的矩阵求得。这使得该变换的复杂度降低了。为了方便变换表达式的书写, 往往利用 \sqrt{N} 对相应的矩阵规格化。例如, $N=8$ 的哈达玛矩阵为

$$\mathbf{H}_8 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \quad (4.41)$$

矩阵右端的值代表该行的列率，即该行中信号符号改变的次数。可见其列率的排列是无规则的。将无序的哈达玛核进行列率排序，即可得到有序的沃尔什变换核：

$$W_8 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} \quad (4.42)$$

二维 WHD 有快速算法，其统计特性与二维 DFT 类似，图像中的直流成分和低序率成分占绝大部分能量，且大部分高序率的变换幅度为零。

4.4 卡胡南-列夫变换 (K-L 变换)

图像通过信道传输时，不可避免会遇到随机噪声，因此我们实际得到的大部分图像是随机图像。卡胡南-列夫变换 (Kahunen-Loeve Transform) 简称为 K-L 变换 (KLT)，也称霍特林变换 (Hotelling) 变换，它以图像的统计特征为基础，是在均方意义下的最佳变换。

设原图像为 X ，采用 KLT 恢复的图像为 \hat{X} ，则 \hat{X} 和原图像 X 具有最小的均方误差 ε ，即

$$\varepsilon = E\{[X - \hat{X}]^T [X - \hat{X}]\} = \min \quad (4.43)$$

设 $N \times N$ 的图像 $f(x, y)$ 在信道中传输了 M 次，则接收到的图像集合为 $\{f_1(x, y), f_2(x, y), \dots, f_i(x, y), \dots, f_M(x, y)\}$ 。对第 i 次获得的图像 $f_i(x, y)$ 可以用 N^2 维向量 X_i 表示：

$$X_i = [f_i(0, 0), f_i(0, 1), \dots, f_i(0, N-1), f_i(1, 0), f_i(1, 1), \dots, f_i(N-1, N-1)]^T \quad (4.44)$$

式中， T 表示转置。将 X_i 视为某个随机向量 X 的一次实现。假设 X 有 M 次实现， X 的数学期望 m_x 可以定义为其估值，即

$$m_x = E\{X\} = \frac{1}{M} \sum_{i=1}^M X_i \quad (4.45)$$

X 的协方差矩阵 C_x 定义为

$$C_x = \frac{1}{M} \sum_{i=1}^M (X_i - m_x)(X_i - m_x)^T = \frac{1}{M} \left[\sum_{i=1}^M X_i X_i^T \right] - m_x m_x^T \quad (4.46)$$

可见， C_x 是一个 $N^2 \times N^2$ 的实对称矩阵。令 λ_i 和 a_i ($i = 1, 2, \dots, N^2$) 分别为 C_x 的第 i 个特征值和特征向量，即

$$A = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{N^2 1} \\ a_{12} & a_{22} & \cdots & a_{N^2 2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1N^2} & a_{2N^2} & \cdots & a_{N^2 N^2} \end{bmatrix} \quad (4.47)$$

根据矩阵论，一个实对称矩阵，其特征向量构成的矩阵是一个正交矩阵，且

$$\mathbf{A}^T \mathbf{C}_x \mathbf{A} = \mathbf{A}^{-1} \mathbf{C}_x \mathbf{A} = \mathbf{A} \quad (4.48)$$

式中, \mathbf{A} 为 \mathbf{C}_x 的特征值构成的对角线矩阵。K-L 变换选取一个上述的正交变换 \mathbf{A} , 使得变换后的图像 \mathbf{Y} 满足

$$\mathbf{Y} = \mathbf{A}(\mathbf{X} - \mathbf{m}_x) \quad (4.49)$$

对于 K-L 变换, 我们可以得出以下结论:

(1) 求图像向量 \mathbf{X} 的 K-L 变换 \mathbf{A} 的问题, 就是求图像协方差矩阵 \mathbf{C}_x 的特征向量的问题。所以, KLT 也称为特征向量变换。

(2) 变换后的新图像 \mathbf{Y} 是对中心化的图像 $(\mathbf{X} - \mathbf{m}_x)$ 进行的正交变换。其数学期望 $\mathbf{m}_y = \mathbf{0}$, 而协方差 $\mathbf{C}_y = \mathbf{A}^T \mathbf{C}_x \mathbf{A} = \mathbf{A}$, 说明 \mathbf{Y} 的元素是各不相关的。KLT 可以通过输入向量 \mathbf{X} 的协方差矩阵 \mathbf{C}_x 的特征向量而得到。

(3) 由 \mathbf{Y} 经逆变换而恢复的原图像 \mathbf{X} 为

$$\mathbf{X} = \mathbf{A}^{-1} \mathbf{Y} + \mathbf{m}_x \quad (4.50)$$

(4) 由于输出变换结果 \mathbf{Y} 的 N^2 个分量彼此是线性无关的, 因此, 如果选取方差最大的前 M 个分量 (主分量) 估计原输入图像 \mathbf{X} 时, 所造成的估值误差将是最小的。因此这种变换又称为主分量变换 (Principal Component Transformation, PCT), 这是主分量分析 (Principal Component Analysis, PCA) 算法的理论依据。相应地, 由于输出误差最小准则等价于取方差最大的成分进行逼近, 所以 KLT 是在最小方差意义下的最佳变换。由变换后的 M 个 y_i 分量恢复图像 \mathbf{X} 时, 估值为

$$\hat{\mathbf{X}} = \mathbf{A}_M \mathbf{Y}_M + \mathbf{m}_x \quad (4.51)$$

式中, \mathbf{A}_M 是 $N^2 \times M$ 维的, \mathbf{Y}_M 是 $M \times 1$ 维的。由于 $M \ll N^2$, 这样复原的图像有效地过滤了随机干扰而成为原图像的最佳逼近, 且有效地压缩了图像数据。

(5) 上述估值误差为

$$\varepsilon = \sum_{i=M+1}^{N^2} \lambda_i \quad (4.52)$$

式中, λ_i 是图像协方差矩阵 \mathbf{C}_x 的特征值。估值误差是 \mathbf{C}_x 的 N^2 个特征值中后 $N^2 - M$ 个较小的特征值之和。由于 λ_i 是按大小递减排列的, 其均方误差将是最小的。

K-L 变换的优点是能够完全去除原信号中的相关性, 因而具有重要的理论意义。其他各种变换可以它为标准来比较性能的优劣。如离散余弦变换具有很强的“能量集中”特性: 能量集中在 DCT 后的低频部分, 而且当信号具有接近马尔可夫过程 (序列中每个元素的条件概率只依赖于它的前一个元素) 的统计特性时, 离散余弦变换的去相关性接近于 K-L 变换。K-L 变换也是图像分析与模式识别中的重要工具, 用于特征抽取, 降低特征数据的维数。例如, 在遥感图像处理中, 一幅 4 波段亮度值构成的灰度图像可以压缩为二维去处理。

K-L 变换的缺点是其基函数取决于待变换图像庞大的协方差矩阵, 因而基函数的形式是不定的, 且变换核是不可分离的, 无相应的快速算法, 故计算量很大。

4.5 二维离散小波变换

小波分析是 20 世纪 80 年代开始逐渐发展成熟的应用数学的一个分支, 其主要特点集中表现在对时间 (二维信号为空间) - 频率的双重分析和多分辨率分析能力, 被誉为“数学显微镜”, 在信号和图像处理等领域具有重要的应用价值。

4.5.1 小波分析的思想来源

傅里叶分析是分析学自 18 世纪逐渐形成的一个重要分支，在长达二百年的发展中一直占主导地位，对近代数学以及物理、工程技术产生了深远的影响。1910 年，Haar（哈尔）提出了 Haar 函数，这是一种正交归一化函数系，以其作为正交规范基的哈尔变换（Haar Transform）收敛均匀而迅速，在图像信息压缩和特征编码等方面获得应用。

Haar 函数的定义域 t 为 $[0,1]$ ，可将它延拓到整个时间轴。Haar 函数表示为 $\text{har}(2^p + n, t)$ ，其中 $p = 1, 2, \dots$ ； $n = 0, 1, \dots, 2^p - 1$ 。

$$\text{har}(0, t) = 1 \quad (4.53a)$$

$$\text{har}(1, t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \end{cases} \quad (4.53b)$$

$$\text{har}(2, t) = \begin{cases} \sqrt{2}, & 0 \leq t < 1/4 \\ -\sqrt{2}, & 1/4 \leq t < 1/2 \\ 0, & 1/2 \leq t < 1 \end{cases} \quad (4.53c)$$

$$\text{har}(3, t) = \begin{cases} 0, & 0 \leq t < 1/2 \\ \sqrt{2}, & 1/2 \leq t < 3/4 \\ -\sqrt{2}, & 3/4 \leq t < 1 \end{cases} \quad (4.53d)$$

Haar 函数系的一般定义为

$$\text{har}(2^p + n, t) = \begin{cases} \sqrt{2^p}, & \frac{n}{2^p} \leq t < \frac{(n+1)/2}{2^p} \\ -\sqrt{2^p}, & \frac{(n+1)/2}{2^p} \leq t < \frac{(n+1)}{2^p} \\ 0, & \text{其他} \end{cases} \quad (4.53e)$$

函数系的前几个函数的波形如图 4.12 所示。这里， $\text{har}(0, t)$ 和 $\text{har}(1, t)$ 为全域函数（global function），而其他函数只有部分区间有数值，为局域函数（local function）。可见，哈尔函数是一种既反映整体又反映局部的正交完备函数系。

哈尔变换具有以下特点：（1）具有尺度和位移两个特性；（2）变换范围窄；（3）其变换特性与图像中的边界或线条的特性十分接近，因此图像中的边缘和线条经哈尔变换后，会产生较大的变换系数，而其他区域的变换系数小。

在信号分析的实际应用中，我们常常需要同时获取频率域和时间域（或空间域）两种信息。如在雷达信号处理中，从回波的到达时间可提供目标的位置信息，而检测回波的频率可知目标的运动速度。这时单纯的傅里叶分析和哈尔变换已无能为力。因为三角函数系中的函数在频率域是局部化的，但在空间域或时间域上却无局部性；相反，Haar 函数系由于缺乏正则性与振荡性，在时间上可以体现局部化，然而它在频率域上的局部性却很差。所以，人们寻求时间-频率域联合分析，能综合三角函数系与 Haar 函数系两者优点的函数来分解非稳态函数。

加窗傅里叶变换（Windowed Fourier Transform, WFT）是符合上述要求的一种时间-频率域联合分析方法，信号 $f(x)$ 的窗口傅里叶变换定义为

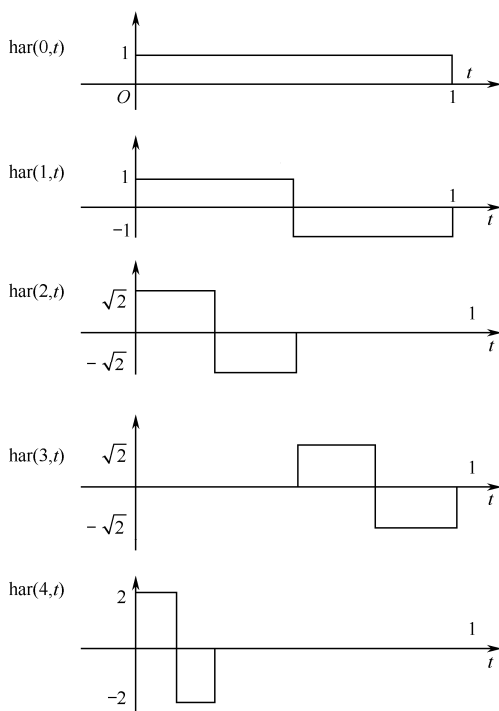


图 4.12 Haar 函数系的前几个函数波形

$$\text{WFT}_f(b, \omega) = \frac{1}{\sqrt{2\pi}} \int_R f(x) W^*(x-b) e^{-j\omega x} dx \quad (4.54)$$

式中, $W(x)$ 是给定的窗函数, $W^*(x)$ 是 $W(x)$ 的共轭对称函数, 一般所选的 $W(x)$ 是关于原点对称的实函数, 即

$$W^*(x) = W(x) = W(-x) \quad (4.55)$$

常见的窗函数具有相对短的时间窗宽, 例如可选为高斯函数, 所以 WFT 也称为短时傅里叶变换 (Short Time Fourier Transform, STFT)。

为了更好地理解 WFT 的意义, 我们将式 (4.54) 改写为

$$\text{WFT}_f(b, \omega_0) = \frac{1}{\sqrt{2\pi}} \int_R f(x) W_{b, \omega_0}^*(x) dx = \langle f, W_{b, \omega_0} \rangle \quad (4.56)$$

式中, $\langle a, b \rangle$ 表示 a 和 b 的内积, 而

$$W_{b, \omega_0}(x) = e^{j\omega_0 x} W(x-b) \quad (4.57)$$

是以函数 $W(x-b)$ 对简谐波 $e^{j\omega_0 x}$ 调制后得到的信号。

如果 $W(x)$ 的中心频率为 ω^* , 时宽为 Δx , 频宽为 $\Delta \omega$, 则时间分辨率取决于窗函数的 Δx , 而频率分辨率取决于 $\Delta \omega$ 。可见, WFT 可以形象地视为以 $(b, \omega^* + \omega_0)$ 为中心、以 $\Delta x \Delta \omega$ 为固定尺寸的时-频矩形窗口对信号进行“细微”观察。

WFT 的重构公式为

$$f(x) = \frac{1}{\sqrt{2\pi}} \iint_{R^2} \text{WFT}_f(b, \omega) W(x-b) e^{j\omega x} d\omega db \quad (4.58)$$

上述的窗口傅里叶变换是一种大小及形状均固定的时频化分析。但实际信号进行时间和频率分析时, 分辨率往往是相对的, 即反映信号高频成分需要较高的时间分辨率, 因此窗函数宽度应该窄一些, 而反映低频成分则需要较高的频率分辨率, 窗函数宽度应该宽一些。显然, 窗口傅里叶变换不能满足上述要求。而小波变换的窗口具有大小 (面积) 固定但形状可改变的特点, 能满足上述时-频局部化分析的要求。

4.5.2 连续小波变换

称按如下方式生成的函数族为连续小波

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \quad (4.59)$$

式中, $a, b \in R, a \neq 0$, a 称为伸缩因子, b 为平移因子, $\psi(x)$ 称为基本小波或母波。

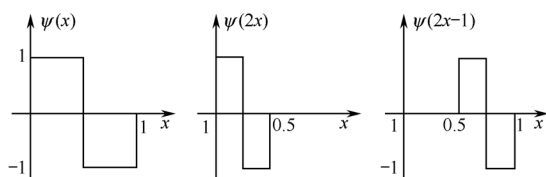


图 4.13 小波函数的平移与扩展

母波可由平移与尺度变换构造小波基函数。如 Haar 函数是一种基本正交基, 也是小波变换中的典型小波, 如图 4.13 所示。其中由 $\psi(x) \rightarrow \psi(2x)$ 为扩展运算, 而由 $\psi(2x) \rightarrow \psi(2x-1)$ 为平移变换。重复使用平移与扩展可以得到下一级的小波函数 $\psi(4x)$ 、 $\psi(4x-1)$ 、 $\psi(4x-2)$ 、 $\psi(4x-3)$ 。这样我们得到 Haar 小波基

$$\{\psi(2^j x - k), j, k \in Z\} \quad (4.60)$$

式中, j 为正整数。设函数 $f \in L^2$, 函数 $\psi(x)$ 满足以下容许性条件

$$C_\psi = \int_R \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (4.61)$$

式中, $\hat{\Psi}(\omega)$ 是 $\psi(x)$ 的傅里叶变换, 则称 $\psi(x)$ 为一容许小波, 并定义信号 $f(x)$ 的连续小波变换的 $W_f(a, b)$ 为

$$W_f(a, b) = \langle f, \psi_{a,b}(x) \rangle = |a|^{-\frac{1}{2}} \int_R f(x) \psi^* \left(\frac{x-b}{a} \right) dx \quad (4.62)$$

由 $W_f(a, b)$ 重构 $f(x)$ 的小波逆变换为

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a, b) \psi_{a,b}(x) \frac{da}{a^2} db \quad (4.63)$$

分析表明, 小波的时间-频率窗使小波变换具有非均匀分布的特点: 对于高的中心频率, 时间-频率窗在时间方向上变窄, 在频率窗方向上变宽, 因此小波有高的时间分辨率; 对于低的中心频率, 小波有高的频率分辨率。

4.5.3 一维离散小波变换

把连续小波变换离散化更有利于实际应用。对 a 和 b 按如下规律取样:

$$a = a_0^m, \quad b = nb_0 a_0^m \quad (4.64)$$

式中, $a_0 > 1$, $b_0 \in R$, $m, n \in Z$, 则由式 (4.60) 得离散小波

$$\psi_{m,n}(x) = a_0^{-\frac{m}{2}} \psi(a_0^{-m}x - nb_0) \quad (4.65)$$

离散小波变换为

$$W_{m,n} = \int_{-\infty}^{\infty} f(x) \psi_{m,n}^*(x) dx = \langle f(x), \psi_{m,n}(x) \rangle \quad (4.66)$$

如果 $f(x)$ 也是离散的, 记为 $f(k)$, 则有

$$W_{m,n} = \sum_k f(k) \psi_{m,n}^*(k) \quad (4.67)$$

小波逆变换的离散形式为

$$f(k) = \sum_{m,n} W_{m,n} \psi_{m,n} = \sum_{m,n} \langle f, \psi_{m,n} \rangle \psi_{m,n} \quad (4.68)$$

S. Mallat 提出的多分辨率分析 (Multi-Resolution Analysis, MRA) 的概念, 在泛函分析的框架下统一了各种具体的小波构造方法, 给出了构造正交小波基的一般方法及其快速算法, 并将小波变换应用于图像分解和重建, 这是小波变换理论上的一个突破性进展。

4.5.4 二维离散小波变换

为了将小波变换应用于图像处理, 需要引入二维变换。二维多分辨率分析与一维情况类似。定义信号 $f(x, y)$ 的连续小波变换 $W_f(a, b_x, b_y)$ 为

$$W_f(a, b_x, b_y) = \langle f, \psi_{a,b_x,b_y}(x, y) \rangle = \frac{1}{a} \int_{R^2} f(x, y) \psi^* \left(\frac{x-b_x}{a}, \frac{y-b_y}{a} \right) dx dy \quad (4.69)$$

式中, b_x 和 b_y 分别表示函数 $f(x, y)$ 在轴上的 x, y 平移量。由 $W_f(a, b_x, b_y)$ 重构 $f(x, y)$ 的小波逆变换为

$$f(x, y) = \frac{1}{a^2} \int_{R^+} \int_{R^2} W_f(a, b_x, b_y) \psi_{a,b} \left(\frac{x-b_x}{a}, \frac{y-b_y}{a} \right) db_x db_y da \quad (4.70)$$

类似地，可以定义二维离散小波变换逼近，并采用 Mallat 二维快速算法求解。与 DFT 类似，可分离二维小波变换最终可转化为两次一维小波变换。

与一维小波变换只用一个尺度函数 $\phi(x)$ 和一个小波函数 $\psi(x)$ 不同的是，二维情形引入的尺度函数换成了 $\phi(x, y)$ 和三个小波函数 $\psi^1(x, y)$ 、 $\psi^2(x, y)$ 和 $\psi^3(x, y)$ 。如果 $\psi(x)$ 是一维多分辨率分析的正交小波基，则二维多分辨率分析的尺度函数和三个方向小波函数为

$$\phi(x, y) = \phi(x)\phi(y) \quad (4.71a)$$

$$\psi^1(x, y) = \phi(x)\psi(y) \quad (4.71b)$$

$$\psi^2(x, y) = \psi(x)\phi(y) \quad (4.71c)$$

$$\psi^3(x, y) = \psi(x)\psi(y) \quad (4.71d)$$

图 4.14 是可分离二维小波变换的频率域分解示意图。作为第一层分解，如图 4.14 (a) 所示， LL_1 表示图像在水平 x 和垂直 y 方向均为低频的分量，即为近似图像（平滑图像）； HL_1 给出了水平 x 方向的高频分量和垂直 y 方向的低频分量，即为水平细节图像； LH_1 给出了水平 x 方向的低频分量和垂直 y 方向的高频分量，即为垂直细节图像； HH_1 给出了 x 和 y 方向的高频分量，即为对角线细节图像。对 LL_1 做第二层分解，如图 4.14 (b) 所示，相应地得到 LL_2 、 HL_2 、 LH_2 和 HH_2 。对 LL_2 可以做类似的处理得到第三层分解，如图 4.14 (c) 所示。这样，将原数字图像做 j 层分解后，可把图像分解成 $3j + 1$ 幅离散图像。通常三层分解就足够精细了。

重构算法按相反的步骤进行，这样就构成了二维小波变换的金字塔结构。

由于小波变换的理论和算法比较复杂，从应用的角度看，读者可以将注意力集中在用 MATLAB 对图像进行小波变换和重构的实现过程中。

对图像进行小波变换的 MATLAB 常用函数有：

(1) 对图像进行一层二维小波分解，常见形式为

$$[CA, CH, CV, CD] = \text{dwt2}(X, 'wname')$$

式中， X 为图像矩阵；'wname' 是使用的小波基函数名称，如可选择双正交样条小波基函数，形式为 biorNr.Nd。

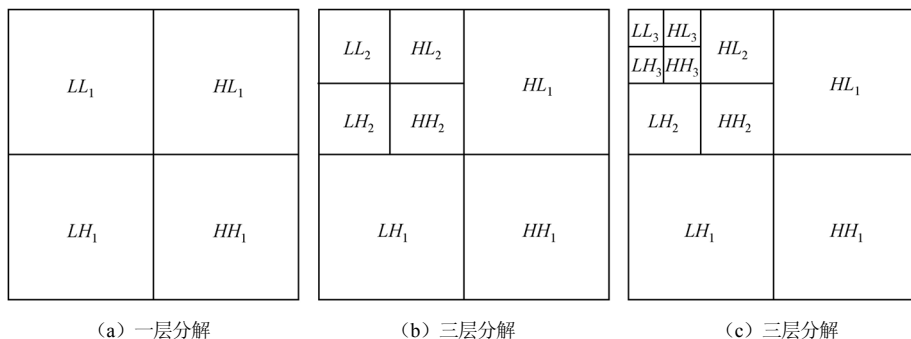


图 4.14 可分离二维小波变换的频率域分解

(2) 查询使用的小波基函数的信息，使用形式为

$$\text{waveinfo}('wname')$$

式中, 小波基名称'wname'可选用'haar' (哈尔小波)、'db' (Daubechies 小波)、'bior' (双正交样条小波) 等。例如, 在命令行状态下键入 `waveinfo('bior')` 进行查询双正交样条小波, 可知 `r` 表示 `reconstruction` (重建), `d` 表示 `decomposition` (分解), `N` 表示相应 FIR 滤波器的阶数; `CA`、`CH`、`CV`、`CD` 分别是输入矩阵 X 小波分解的近似系数矩阵、水平细节系数、垂直细节系数和对角线细节系数。

(3) 对二维小波分解的图像进行各种分量的重构, 常见函数形式为

$$Y = \text{upcoef2}(O, X, \text{'wname'}, N)$$

式中, X 是分解后的细节信号, Y 是重构的细节信号分量; N 表示对矩阵 X 的系数进行重建的步骤数, 即重构的层数, 默认值为 1。O 是细节信号的类型。如果 $O = 'a'$, 则表示对信号的近似系数进行重建; 否则, 如果 $O = 'h'$ 、 $'v'$ 或 $'d'$, 则分别对水平、垂直或对角线细节进行重建。

(4) 对应上述的一层二维小波变换 `DWT2` 函数, 进行一层二维小波变换逆变换, 常见形式为

$$X = \text{idwt2}(CA, CH, CV, CD, \text{'wname'})$$

`idwt2` 函数采用'wname'所指示的小波、已重建的基于近似矩阵 `CA`, 以及水平细节 `CH`、垂直细节 `CV` 和对角线细节 `CD` 计算原图像矩阵 X 。

(5) 对重构的图像进行量化编码, 常见函数形式为

$$Y = \text{wcodemat}(X, \text{NBCODES}, \text{OPT}, \text{ABSOL})$$

式中, X 为待进行量化编码的矩阵, Y 为编码矩阵。在编码中, 把矩阵 X 中元素绝对值最大的作为 `NBCODES` (整数), 绝对值最小的作为 1, 其他元素依其绝对值的大小在 1 与 `NBCODES` 中排列。当 `OPT` 为'row'时, 做行编码; 当 `OPT` 为'col'时, 做列编码; 当 `OPT` 为'mat'时, 做全局编码, 即把整个矩阵中元素绝对值最大的元素作为 `NBCODES`, 最小的作为 1。当 `ABSOL` 为 0 时, 该函数返回输入矩阵 X 的一个编码版本, 当 `ABSOL` 非 0 时, 返回 X 的绝对值。

【例 4.6】 `bior3.7` 是双正交样条小波对应的滤波器。应用 MATLAB 对 `wbarb.mat` 图像实现小波变换的例子。

【解】 MATLAB 程序如下:

```
load wbarb;%           %从磁盘调入磁盘文件 wbarb.mat
image(X);              %将矩阵 X 显示为图像
colormap(map);         %配合函数 image() 画出连续的灰度图

%对 X 进行二维小波变换, bior3.7 是双正交样条小波对应的滤波器
[cA1, cH1, cV1, cD1] = dwt2(X, 'bior3.7');
A1 = upcoef2('a', cA1, 'bior3.7', 1);
H1 = upcoef2('h', cH1, 'bior3.7', 1);
V1 = upcoef2('v', cV1, 'bior3.7', 1);
D1 = upcoef2('d', cD1, 'bior3.7', 1);

figure; colormap(map);
nbcol = size(map, 1);
cod_X = wcodemat(X, nbcol);
cod_cA1 = wcodemat(cA1, nbcol);
subplot(2, 2, 1); image(cod_cA1);
title('Approximation A1')
cod_cH1 = wcodemat(cH1, nbcol);
subplot(2, 2, 2); image(cod_cH1);
title('Horizontal Detail H1')
cod_cV1 = wcodemat(cV1, nbcol);
```

```

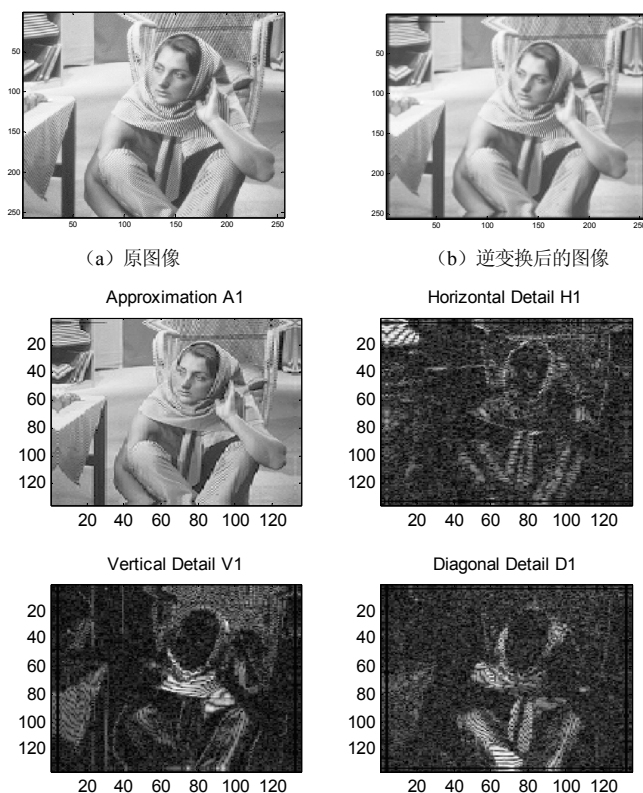
subplot(2,2,3); image(cod_cV1);
title('Vertical Detail V1')
cod_cD1 = wcodemat(cD1,nbcol);
subplot(2,2,4); image(cod_cD1);
title('Diagonal Detail D1')

% 显示逆变换图像，为了进行比较，调整其尺寸和灰度范围与 x 一致
Y = IDWT2(A1,H1,V1,D1, 'bior3.7');
Y1 = imresize(Y,0.461);
Y2 = min(min(X))+(max(max(X))-min(min(X)))/...
    (max(max(Y1))-min(min(Y1)))*(Y1-min(min(Y1)))

% 显示已编码的图像
figure; image(Y2);colormap(map);
dec2d = [cod_cA1, cod_cH1; cod_cV1, cod_cD1];
figure; image(dec2d); colormap(map);

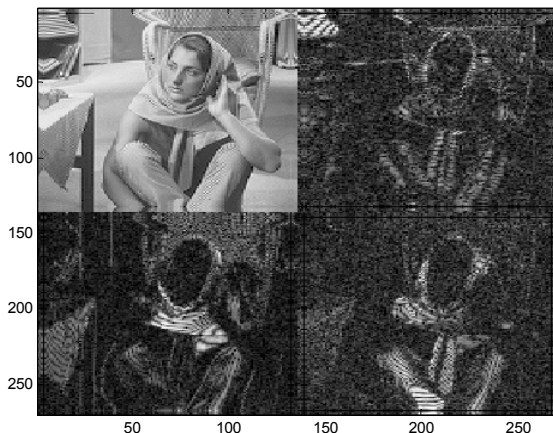
```

图 4.15 (a) 是原图像 `wbarb.mat`，图 4.15 (c) 是原图像的一阶小波变换的结果，分别是低频分量和 3 个方向上的高频分量，选择 `NBCODES` 参数使近似图像与图像亮度大致一致、突出显示细节图像。图 4.15 (b) 是小波逆变换后的图像，为了与原图像进行比较，对逆变换后图像的灰度范围和尺寸均进行了修正。



(c) 分别显示一层小波变换的 4 个分量

图 4.15 一层小波变换



(d) 同时显示一层小波变换的4个分量

图 4.15 一层小波变换 (续)

4.6 二维数字滤波器

与一维滤波器广泛用于一维信号处理类似, 二维线性滤波器在图像处理中也发挥重要作用。如图 4.16 所示, 设尺寸为 $M \times N$ 的待处理输入为 $f(x, y)$, 其 DFT 为 $F(u, v)$; 系统的单位脉冲响应为 $h(x, y)$, 其 DFT 为 $H(u, v)$; 已处理的输出图像为 $g(x, y)$, 其 DFT 为 $G(u, v)$, 则

$$g(x, y) = f(x, y) * h(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} h(m, n) f(x-m, y-n) \quad (4.75)$$

$$G(u, v) = F(u, v) \cdot H(u, v) \quad (4.76)$$

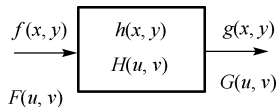


图 4.16 二维数字滤波器空域和频域框图

显然, 式 (4.75) 表示了图像空域处理的输入输出卷积关系, 式 (4.76) 表示了图像频域的输入输出乘积关系。空域法将在图像增强等章节详细介绍, 这里主要强调应用频域进行处理的特点。采用频域进行处理不仅可以根据图像的频谱进行必要的滤波, 而且还便于某些特征的提取。例如, 一般图像的对比度和动态范围取决于图像的低频分量, 而边缘轮廓和局部细节取决于图像的高频部分。一定频率的周期性的干扰会对图像画面产生严重影响。采用低通滤波器、高通滤波器、带通滤波器或带阻滤波器, 即可完全或部分达到我们对图像质量的需求。图像从空域变换到频域, 将显著地降低相邻像素等信息的相关性, 有利于图像数据量的压缩。

二维数字滤波器的设计方法与一维滤波器的设计方法类似, 如二维 FIR 滤波器采用窗口设计法(用一维、二维窗口方法设计的 MATLAB 函数分别为 fwind1、fwind2)、频率采样法设计法 (MATLAB 函数为 fsamp2)。频率转换设计法 (MATLAB 函数调用格式为 $H = \text{ftrans}(B, T)$) 采用频率变换 (如 McClellan 变换: $T = [1 \ 2 \ 1; 2 \ -4 \ 2; 1 \ 2 \ 1]/8$) 生成对应于偶对称、奇数长度的一维 FIR 滤波器 B 的二维 FIR 滤波器 H。二维数字滤波器的频率特性 (MATLAB 函数为 freqz2) 不难从一维数字滤波器的频率特性增加一维频率变量获得。

一个二维 FIR 低通数字滤波的实例如图 4.17 所示。该 21 阶滤波器采用二维 Blackman 窗口设计, 对应于 (u, v) 的频域 (ω_1, ω_2) 截止频率均为 $\pi/2$ 。原始测试图像为 rectangles, 左上方为图像的水平 and 垂直方向均为高频细节, 而右下方则相反。右上方和左下方分别表现图像水平低频、垂直高频和水平高频、垂直低频两种情况。可见, 滤波后的图像在细节处产生了明显的模糊现象, 符合低通特性。

与图 4.17 对应的一个二维 FIR 高通数字滤波的实例如图 4.18 所示，可见滤波后的图像在慢变化部分产生了明显的模糊现象，符合高通特性。

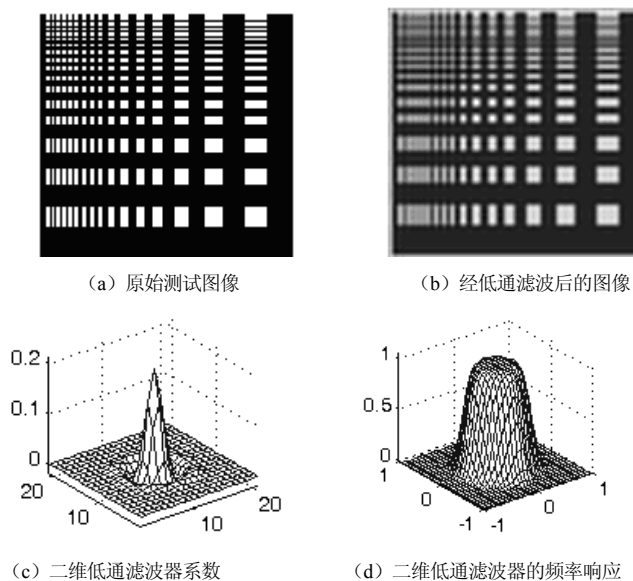


图 4.17 二维 FIR 低通数字滤波

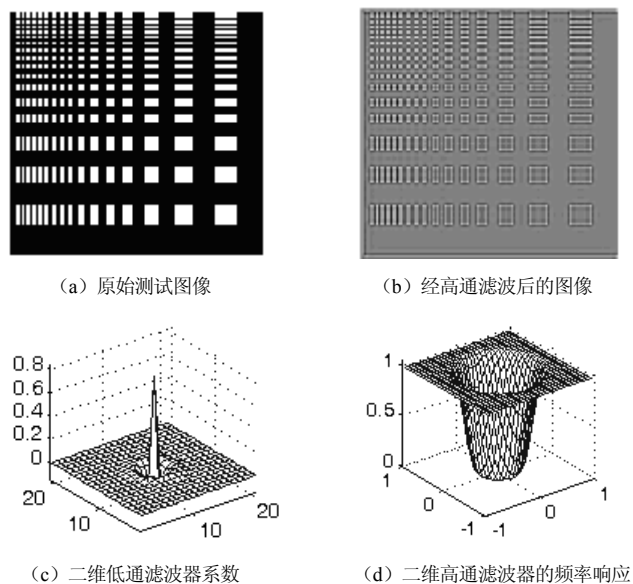


图 4.18 二维 FIR 高通数字滤波

4.7 实验：图像变换与二维数字滤波

1. 实验目的

- (1) 了解图像正变换和逆变换的原理。
- (2) 理解图像变换系数的特点。

- (3) 掌握常用图像变换的实现过程。
- (4) 掌握图像的频谱分析方法。
- (5) 了解图像变换在图像数据压缩等方面的应用。
- (6) 掌握二维数字滤波器的作用以及在图像滤波中的实际应用。

2. 实验主要仪器设备

- (1) 台式或笔记本电脑。
- (2) 安装了图像处理工具箱的 MATLAB 软件。

3. 实验原理

- (1) 二维离散傅里叶变换、余弦变换、小波变换的正逆变换公式。
- (2) 图像的频谱分析原理。
- (3) 讨论正交变换的应用。
- (4) 二维 FIR 滤波器在图像滤波中的应用。

4. 实验内容

(1) 在 MATLAB 环境中, 进行图像的离散傅里叶变换和离散余弦变换, 观察图像的频谱并减少 DCT 系数, 观察重建信号和误差信号, 理解正交变换在压缩编码中的应用。

(2) 在 MATLAB 环境中, 进行图像的离散小波变换, 观察图像的近似图像和各方向的细节图像, 观察重建图像, 理解小波变换在图像特征检测 (如边缘检测、方向检测等) 中的应用。

(3) 在 MATLAB 环境中, 实现图像的二维 FIR 滤波。

5. 实验步骤

- (1) 选择典型图像作为研究对象。
- (2) 显示原始图像。
- (3) 进行图像变换。
- (4) 对图像进行处理 (如选择不同个数的变换系数可以进行压缩、选择不同方向的频谱可以进行特征检测等)。
- (5) 对图像进行逆变换复原图像, 观察重建图像和误差图像并进行分析。
- (6) 实现图像的二维 FIR 滤波, 分别采用窗口设计法、频率采样设计法和频率变换法。注意滤波后的图像和原始图像的差异并分析原因。

6. 实验报告要求

- (1) 说明实验采用的正交变换的原理。
- (2) 分析重建图像、误差图像和变换系数的关系。
- (3) 通过实验结果说明三种变换的优点以及各自适用的应用场合。
- (4) 分析图像的二维 FIR 滤波的结果和各种设计法的优缺点。

7. 预习要求

掌握主要图像正交变换和二维滤波的基本原理, 了解 MATLAB 中相关正交变换和二维滤波函数的使用方法。

8. 思考题

- (1) 信号从一维变换到二维变换, 在计算方法上有何新的特点? 如何实现快速算法?
- (2) 如何对重建图像进行主观评价和客观评价, 分析重建图像的质量与重建系数的关系?
- (3) 对图像单方向的滤波(行滤波、列滤波)与双方向滤波的计算复杂度进行比较, 采用快速算法呢?

本章小结

图像处理可以直接在空间域进行, 也可以通过变换的手段达到更好的效果。因此, 图像变换是图像处理中常用的有效分析手段, 对实现某些图像处理和后期的图像分析起着十分重要的作用。空间域信息是视觉最容易接受的信息, 但频率域或小波域提供了视觉不易直接感受的信息; 从图像的空间域信息变换为频率域或小波域信息可能使某些图像处理问题简化; 频率域或小波域信息有利于某些图像特征的提取。

为了达到图像处理、分析、识别、传输和存储等目的, 图像变换一般为正交变换。正交变换可以显著地减少图像数据的相关性, 可以实现用较少的数据量表示原始图像及其特征。

二维数字滤波是从系统角度, 对图像进行处理的重要手段。从一维滤波器的设计方法推导出二维滤波器的设计方法, 重要的是如何针对具体的图像和处理的目的选择合适的二维滤波器和参数, 读者要结合实例善于分析和应用。

本章针对图像处理中常用的变换进行了详细的介绍, 特别分析了小波变换的思想和实现方法。本章的重点是傅里叶变换、离散余弦变换、二维滤波, 难点是小波变换。读者可以结合实例来理解各个变换和二维滤波的意义。

思考题与习题

- 4.1 二维傅里叶变换的可分离性有什么实际意义?
- 4.2 图像处理中正交变换的目的是什么? 图像变换主要用于哪些方面?

4.3 已知 $f(x,y)=\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, 按 x 方向进行 FFT, 得到列变换 $F(u,y)=$

$\begin{bmatrix} 2 & 2 & 0 & 0 \\ 1-j & 1-j & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1+j & 1+j & 0 & 0 \end{bmatrix}$, 为什么 $f(x,y)$ 的前 2 行数值一样, 但 $F(u,y)$ 的前 2 行却不一样?

4.4 证明 $f(x,y)$ 的直流成分为 $F(0,0)=\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)$, 并说明其物理意义。

4.5 由 DFT 的定义直接证明 DFT 的共轭对称性。

4.6 本章的例 2 中, 在求解图像频谱的程序中为了增强显示效果, 用对数对 DFT 的幅度进行压缩, 然后将频谱幅度的对数值用在 0~10 之间的值进行显示。研究: (1) 对上述结果与不进行上述处理的直接显示效果进行比较。(2) 对将频谱幅度的对数值用不同范围的值进行显示的结果进行比较。对研究的结果做出结论。

4.7 对 0 到 $2N-1$ 的 $2N$ 个点的离散周期序列 $f_c(n)$ 做 DFT, 设 $W_{2N} = e^{-j\frac{2\pi}{2N}}$, 证明 $F_c(k) =$

$$\sum_{n=0}^{2N-1} f_c(n) W_{2N}^{nk} = W_{2N}^{\frac{k}{2}} \sum_{n=0}^{N-1} f(n) \cos\left[\frac{\pi(2n+1)k}{2N}\right].$$

4.8 利用 MATLAB 函数 `dct2` 对一幅 8×8 的图像进行 DCT 变换和反变换, 进行原始图像和重建图像的误差比较分析。变换的参考图像为:

$$I = \begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 65 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 58 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 65 & 78 & 94 \end{bmatrix}.$$

采用 Z 扫描方式 (参见第 5 章 5.15) 保留 20 个 DCT 变换系数进行重构图像, 即将第 21 个以后的变换系数置为 0。计算重建图像 J 与原始图像 I 的均方误差, 比较两者的差异。如果保留横向与纵向的坐标之和小于 8 的系数, 比较上述结果。如果保留所有变换系数, 再比较上述结果。这时有误差存在吗? 如果有误差, 出现在何处?

4.9 在沃尔什变换中, 当 $N=8$ 时, 如果 $z=6$, 求 $b_i(z)$ 。

4.10 离散的沃尔什变换与哈达玛变换之间有哪些异同?

4.11 求 $N=8$ 对应的沃尔什变换核矩阵。

4.12 已知二维数字图像矩阵为 $f = \begin{bmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{bmatrix}$, 求此图像的二维沃尔什变换。

4.13 小波基函数和傅里叶变换基函数有何区别?

4.14 一维小波变换如何同时实现时间-频率分析功能? 如何扩展到在二维图像的空间-频率分析?

4.15 采用窗口设计法对 `lena` 图像进行低通滤波, 归一化截至频率为 $(0.5, 0.5)$, 滤波器系数为 10 阶。

4.16 采用频率采样设计法对 `lena` 图像进行高通滤波, 归一化截至频率为 $(0.5, 0.5)$, 滤波器系数为 10 阶。

4.17 采用频率变换法对 `lena` 图像进行二维带通变换, 归一化频率截至频率分别 $(0.4, 0.6)$, 滤波器系数为 10 阶。

第5章 图像编码与压缩

内容提要

本章主要介绍图像信号压缩原理、典型的图像压缩编码算法、图像编码器/解码器的实现、图像压缩标准等。图像压缩编码主要有统计编码、预测编码、变换编码、混合编码及一些新型的压缩算法等。编码的图像类型有二值图像、灰度图像和彩色图像。图像编码压缩的国际标准有 JPEG、JBIG、JPEG2000 等。本章主要介绍灰度图像的编码与压缩原理和方法，在第 10 章将介绍彩色图像的编码与压缩问题。

知识要点

- 信息论中的有关概念：信息，信息量，信息熵，冗余度。
- 统计编码。
- 预测编码。
- 变换编码。
- 混合编码。
- 现代压缩编码方法有分形编码、模型基编码等。
- 静态图像压缩标准：JPEG、JBIG、JPEG2000 等。

教学建议

- 本章的教学安排建议用 6 学时讲授。
- 先修知识主要有信息论与编码、数字信号处理（正交变换）。本章的内容又是后续课程（如图像通信、现代声像技术、多媒体技术等）的先修知识。
- 教学基本要求：理解图像数据压缩的必要性及其理论基础，了解 JPEG、JPEG2000 标准和编码器/解码器的软件与硬件实现。建议将小波变换、分形编码压缩等现代图像压缩方法、JPEG2000 国际标准作为选讲内容。

5.1 概述

数据编码的目的各异,如信息保密、信息的压缩存储和传输等。图像数据是一种十分重要且数据量很大的信息源,特别是多媒体、物联网、大数据及云计算等技术兴起后,它已成为现代信息构成最重要的组成部分,其编码压缩技术受到人们的普遍关注。人们通过数码相机、扫描仪等数字图像摄取设备获得大量照片、图片等静态图像信息并能够永久保存。在图像通信、多媒体网络通信等场合,压缩编码已形成一系列静态图像和视频国际标准,它们为信息高速公路的铺设奠定了技术基础。4G 移动通信系统能够以 100 Mbps 的速度下载,上传的速度也能达到 20 Mbit/s。即便有如此宽的带宽,系统传输的信息大都还是压缩后的视频信息。星际通信的数据传输带宽更是极其宝贵的。例如,“好奇号”火星车主照相机硬件 MPEG-2 视频压缩(I 帧: 2 bpp(比特/像素),P 帧: 0.67 bpp),与地球的直接数据带宽大约 8 kbit/s,但与火星“奥德赛号”探测卫星的最理想传输带宽则能达到 2 Mbit/s,而“奥德赛号”与地球的通信带宽为 256 kbit/s。当探测器从漫游车上空飞过时,每次能通信 8 分钟,最多能传输 250 Mbit 的数据,而这些数据需要花 20 小时才能传输到地球。如果不压缩,其传输效率更不可想象。本章主要介绍静态图像压缩编码的原理、应用及有关的国际标准,而视频图像的编码、传输和存储是现代声像技术、多媒体技术等课程研究的重要内容。

5.1.1 数据压缩的基本概念

数据压缩就是以较少的数据量表示信源以原始形式所代表的信息,目的在于节省存储空间、传输时间信号频带或发送能量等。这些概念无论是针对静态的文字、图像,还是针对动态的音频、视频都是适用的。各种压缩方法不尽相同,但系统组成可由图 5.1 所示。

由图 5.1 可知,数据压缩处理是由编码和解码两个过程组成的。编码对原始的信源数据进行压缩,便于传输或存储;解码是编码的反过程,它将不能被用户(信宿)直接使用的数据还原成可用数据。在图 5.1 中,信源编码主要解决压缩的有效性问题,而信道编码主要解决编码的可靠性问题。从原理上看,压缩主要靠前者,而后者是压缩过程能够可靠实现的保证。本章的讨论基于信源编码。

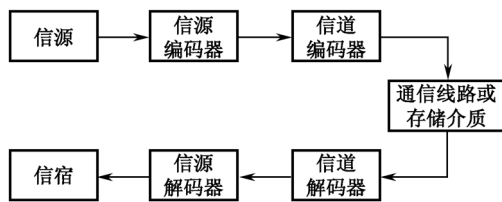


图 5.1 数据压缩系统组成图

信息论认为:以熵(entropy)代表信源所含的平均信息量,若信源编码的熵大于信源的实际熵,则信源中的数据一定存在冗余度。冗余数据的去除不会减少信息量。信息量与数据量的关系可由式(5.1)给出

$$I = D - \text{du} \quad (5.1)$$

式中, I 、 D 、 du 分别表示信息量、数据量和冗余量。自然,冗余量是可以压缩的。在实际应用场合,压缩过程应尽量保证去除冗余量而不会或较少减少信息量,即压缩后的数据要能够完全或在一定的容差内近似恢复。完全恢复被压缩信源信息的方法称为无损压缩方法或无失真压缩方法,而近似恢复的方法称为有损压缩方法或有失真压缩方法。显然,对于有损压缩来说,采用同一压缩方法对同样的信源进行压缩,压缩程度越高,信息损失越大。所以,人们常常在压缩程度和保真度之间权衡。

有关数据压缩的研究理论和实用系统日臻完善。从 PCM 编码理论开始到现在的多媒体数据压缩标准 JPEG、MPEG,发展速度异常迅速,出现了大量实现有关算法的大规模集成电路和实用软件。

5.1.2 图像编码压缩的必要性

采用数字技术会使信号处理的性能大为提高，但其数据量的增加也是十分惊人的。图像数据更是多媒体、网络通信等技术重点研究的压缩对象。不加压缩的图像数据是计算机的处理速度、通信信道的容量等所无法承受的。

图像信号的数据量可以用式（5.2）表示：

$$V = whd/8 \quad (5.2)$$

式中， V 、 w 、 h 和 d 分别表示图像数据量（字节，byte，B）、图像宽度（像素数，pel）、图像高度（像素数，pel）和图像深度（位，bit）。图像的尺寸为 wh 。

对于 $640 \times 480 \times 256$ 色的中等质量静态图像， $w = 640$ pel， $h = 480$ pel， $d = 8$ bit，所以 $V = 300$ KB。高清晰度电视（HDTV）的数码率可达400 MB/s以上。这样的数据率是与当前信息存储介质的容量、计算机的总线速度及网络的传输率不相匹配的，同时大量的数据还会带来较大的安全风险。尽管人们在存储介质、总线结构和网络性能等方面不断有新的突破，但数据量的增长速度远超过硬件设施的提高水平，以上的矛盾仍然无法缓解。

如果将上述图像信号压缩几倍、十几倍甚至上百倍，将十分有利于图像的传输和存储。可见，在现有硬件设施条件下，对图像信号本身进行压缩是解决上述矛盾的主要出路。

5.1.3 图像编码压缩的可能性

图像数据量大，同时冗余数据也是客观存在的。在有些图像中，可压缩的可能性很大。一般图像中存在着以下数据冗余因素。

1. 编码冗余

经典的数据压缩理论建立在信息论基础之上。来自于信息论的式（5.1）已明确指出，去除信源编码中的冗余量可以在对信息无损的前提下减少代表信息的数据量。对图像进行编码时，要建立表达图像信息的一系列符号的码本（codebook）。如果码本不能使每个像素所需的平均比特数最小，则说明存在编码冗余，就存在压缩的可能性。

2. 像素间的相关性形成的冗余

图像由按一定规则排列起来的像素组成。所以，在同一扫描行的邻近像素间、在同一帧的邻近行间、在活动图像中的同一位置的相邻帧像素间的灰度和色度往往相同或相近，我们称这种相关性为像素间冗余或空间冗余。例如，二值传真信息的数据流是由大量的“白”数据串和“黑”数据串构成的。我们没有必要将全部数据加以发送，只需以“起始串，第二串长，第三串长……”表示即可。利用像素间的相关性进行压缩的原理，实际上是将二维像素的矩阵表示成数据量更小的表达形式。

3. 视觉特性和显示设备引起的冗余

由于受生理和心理的影响，人眼对黑白与彩色信息的空间分辨率、时间分辨率等是有限的。光路上不一致的图像在视觉上可能是一样的。人眼对空间细节、灰度层次、时间变化等分辨力的极限值是分别测量出来的，各自都可能比较高，但是这些分辨力不会同时达到极限值。例如，人眼对空间（时间）变化大的区域（时刻）的灰度分辨力就很差。因此，在图像的细节部分或跳变的边沿，灰度的误差是不容易被发现的，只需分配少许的量化比特。同样，在时间上迅速变化的运动图像也是无须顾及灰度误差的。当然，人眼对静止图像的分辨力较高，平缓的区域需要精细地量化，但这时周围像素的相关性又很大，又可以利用上述的相关性对冗余加以压缩。

同样,显示设备由于受到制作工艺的影响,有时也不能完全反映图像的每个细节。只要满足视觉和显示的要求,一定限度的失真允许的。这给有损压缩显著地提高压缩比带来了可能。事实上,与发射台的信号相比较,人们从电视上看到的都是有一定程度上失真的图像,但这并没有影响人们对电视内容的欣赏和理解。因此,在许多压缩系统中,主观评价也是衡量压缩特性的重要技术指标。

5.1.4 图像编码压缩的技术指标

一般地,图像压缩应能做到压缩比大、算法简单、易于用硬件和软件实现、压缩和解压缩实时性好、解压缩恢复的图像失真小等。但这些指标对同一压缩方法很难统一,在实际系统中往往需要抓住主要矛盾,全面权衡。下面介绍常用的图像压缩技术指标。

1. 图像熵与平均码长

熵是指信源的平均信息量。如果一幅图像像素的灰度级为 x_i ($i = 1, 2, \dots, M$), 若 x_i 出现的概率为 $p(x_i)$, 则图像熵定义为

$$H(x) = -\sum_{i=1}^M p(x_i) \log_2 p(x_i) \quad (5.3)$$

图像熵表示像素灰度级集合的平均比特数。熵的单位为比特/字符。熵具有以下性质:

- (1) 当 M 级灰度出现的概率相等时, 即 $p(x_i) = 1/M$ 时, 有最大熵值 $H(x) = \log_2 M$ 。
- (2) 在极端情况下, 当 $p(x) = 0$ 或 $p(x) = 1$ 时, $H(x) = 0$ 。这表明确定性信号的熵值为 0。
- (3) 随机性图像信号的熵非负, 满足 $0 < H(x) \leq \log_2 M$ 。
- (4) 令 $M = 2^L$, 则在各灰度等概率的情况下, $p(x_i) = 2^{-L}$, $H(x) = L$ 。在不等概率的情况下, $H(x) < L$ 。平均码字长度简称平均码长, 是码字长度的数学期望, 即

$$R(x) = \sum_{i=1}^M p(x_i) \cdot L_i \quad (5.4)$$

显然, $R(x)$ 的单位也是比特/字符。根据香农的信息保持编码定理, 要保持信源的全部信息, 就必须有

$$R(x) \geq H(x) \quad (5.5)$$

否则一定会产生信源的编译码失真。

2. 图像冗余度与编码效率

从上面的分析可以看出, 图像无损压缩的码长以 $H(x)$ 作为极限。因此可以定义冗余度 r 为

$$r = 1 - \frac{H(x)}{R(x)} \quad (5.6)$$

将编码效率 η 定义为

$$\eta = \frac{H(x)}{R(x)} = \frac{1}{1+r} \quad (5.7)$$

可见, 当平均码长接近 $H(x)$ 时, 冗余度下降至 0, 编码效率提高至 1。这正是高效编码追求的目标。

3. 压缩比

压缩比 C_r (compression ratio) 是衡量数据压缩方法的压缩程度的一个指标, 反映了压缩效率。通常将 C_r 定义为压缩前图像每像素码长的平均码长与压缩后每像素码长的平均码长之比, 即

$$C_r = \frac{\sum_{x=1}^{M-1} \sum_{y=1}^{N-1} r_b(x, y)}{\sum_{x=1}^{M-1} \sum_{y=1}^{N-1} r_c(x, y)} = \frac{\bar{r}_b}{\bar{r}_c} \quad (5.8)$$

式中，图像的尺寸为 $M \times N$ ， r_b 为原图像像素使用的码长， r_c 为压缩后的图像像素使用的码长， \bar{r}_b 为原图像像素使用的平均码长， \bar{r}_c 为压缩后每像素使用的平均码长。若 $C_r > 1$ ，则 C_r 值越大，压缩效率越高。

4. 客观保真度评价

客观保真度准则指原图像 $f(x, y)$ 和压缩后的恢复图像 $g(x, y)$ 之间的均方误差、信噪比和峰值信噪比等的度量。

(1) 均方误差 MSE

均方误差 (Mean Square Error) 为

$$SME = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - g(x, y)]^2 \quad (5.9)$$

(2) 信噪比 SNR

客观评价 SNR (Signal to Noise Ratio) 指压缩前的图像信号 $f(x, y)$ 方差 σ_x^2 与解压缩后重建图像 $g(x, y)$ 误差方差 σ_e^2 的比值。定义如下：

$$SNR = 10 \log_{10} \left(\frac{\sigma_x^2}{\sigma_e^2} \right) \quad (5.10)$$

显然，SNR 越大，在压缩过程中引入的失真越小，图像质量越高。

(3) 峰值信噪比 PSNR

有损压缩图像质量或失真程度一般用峰值信噪比 PSNR (Peak Signal to Noise Ratio) 指标来衡量。假设 $f(x, y)$ 方差 σ_x^2 与解压缩后重建图像 $g(x, y)$ 误差方差 σ_e^2 的比值。定义如下：

$$PSNR = 10 \log_{10} \left(\frac{A^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x, y) - g(x, y)]^2} \right) \quad (5.11)$$

式中， A 为重建图像图像的最大值。

5. 主观评价

客观评价便于计算，是目前常用的量化指标，但是这些指标不能完全反映人类视觉效果。除了机器视觉，许多图像是为人类服务的，信宿实际上是人的眼睛，当然，最终的评价标准是人的主观感觉。主观评价结果可以用参与测试组全体组员的平均判分 MOS (Mean Opinion Score) 来统一衡量。表 5.1 列出了电视等系统中图像质量的主观评价等级及有关说明。

表 5.1 图像质量的主观评价等级

评 分	评 价	说 明
5	优秀	图像质量非常好
4	良好	图像质量高，有很小的干扰但不影响观看
3	中等	图像质量可接受，但有一些干扰，对观看稍有妨碍
2	差	图像质量差，对观看有妨碍
1	很差，劣	图像质量很差，无法观看

图像的主客观两种评价之间存在着密切的联系。但一般来说,客观评价高的主观评价也高,因此在图像编码的质量评价时,首先做客观评价,以主观评价为参考。

5.1.5 数据压缩方法的分类

数据压缩的方法很多,而且人们还在不断地研究新方法。一般数据压缩按信息损失的程度来分类。常见的有 20 多种常用数据压缩方法(见图 5.2),且这些方法在图像压缩中均有应用。本章将对其中的重要方法进行介绍。

在无损压缩(lossless compression)中,Huffman 编码和 Shannon 编码根据概率分布特性确定码长;行程编码根据连续灰度的行程来确定编码;算术编码随信源数据不断缩小实数区间,然后用一个与实数对应的二进制码代表被编码的信息;轮廓编码根据相同灰度的区域边界线进行编码。

在有损压缩(lossy compression)中,预测编码根据相邻像素的相关性来确定后继像素的预测值,若用差值进行编码,则可以压缩数据量;变换编码对原始图像进行正交变换,在变换域进行抽样达到压缩的目的;混合编码将两种编码方法结合起来,如将预测编码与变换编码相结合,以取得更好的效果。

在现代压缩编码方法中,分形编码利用宏观与微观的相似性来压缩数据量,可以获得极大的压缩比。该方法压缩过程的计算量较大,但解压缩很快,适用于图像数据的存储和重现。

模型基(model-based)编码也是一种新型压缩方法。该方法在发送端利用已知且变化慢的场景得到数据量不大的模型参数,在接收端利用综合模型参数恢复原始图像。这一编码方法对于实时实现电视会议等具有显著意义。

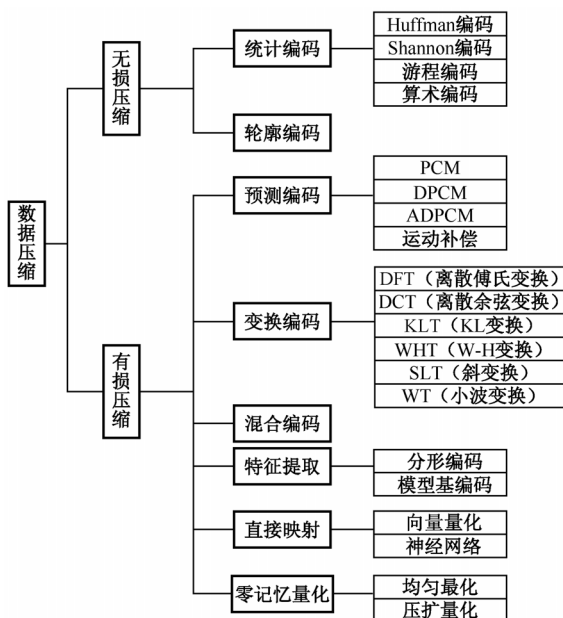


图 5.2 常见数据压缩技术的分类

5.2 统计编码

统计编码(statistic coding)是根据信源的概率分布特性,分配具有唯一可译性的可变长码字(variable-length code),降低平均码字长度,以提高信息的传输速度,节省存储空间。统计编码的基本原理是,在信号概率分布情况已知的基础上,概率大的信号对应的码字短,概率小的信号对应的码字长,这样就降低了平均码字长度。各种统计编码的差异在于信号与编码对应的规则不同,性能亦有差异。1843 年, Morse(莫尔斯)电报码就是最初的变长码数据压缩的例子,它把出现概率最大的字母“e”设置为最短的一个“点”。概率分布特性越不均匀,统计编码越有用武之地。

5.2.1 Huffman 编码

早于 Huffman 编码的提出,C. E. Shannon 于 1948 年提出了将信源符号依其概率降序排列,用符号序列累积概率的二进制表示作为对信源的唯一可译编码。在灰度级有限的情况下,Fano 编码的平均码长较 Shannon 编码为短,但它不是唯一可译编码。当灰度级较多时,Shannon 编码的平均码长与 Fano

编码的平均码长趋向于一致。D. A. Huffman（哈夫曼）在 1952 年根据可变长最佳编码定理，提出了依据信源集中符号出现的概率分配不同长度的唯一可译码的算法。接收端在得到哈夫曼编码后，通过解码可以得到与输入完全一致的信号。

可变码长最佳编码定理 在变长编码中，如果码字长度严格按照信号中符号出现概率大小的相反顺序排列，则平均码字长度一定小于其他符号顺序排列方式的平均码字长度。

1. 前缀码（prefix code）

一组唯一可译码中的任意一个码字都只与一种信号存在对应关系。为了译码的需要，在唯一可译码中的前缀码保证任意一个码字都不是其他码字的前缀。例如，有一维图像的符号集合为 $f(i) = \{f(1), f(2), f(3), f(4)\}$ ，设定的码字集合 $c(i) = \{0, 10, 110, 111\}$ 。编码系统解码时，只要一遇到“0”，便知道对应的是 $f(1)$ 。若接收到的是“1”，则等待下一个比特，若下一个比特为“0”，即确定是 $f(2)$ ，若下一个比特是“1”，则等待第三个比特。若第三个比特为“1”，则可判定信号为 $f(4)$ ，否则为 $f(3)$ 。若一前缀码为 010111100，则译码的输出信号为序列 $\{f(1), f(2), f(4), f(3), f(1)\}$ 。可见前缀码保证了这样译出的码字具有唯一性和“即时性”。

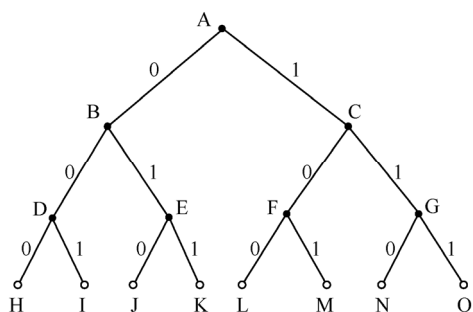


图 5.3 前缀码的构造

前缀码的构造常采用根节点的概率为 1 的二进制码树。这是典型的二叉树数据结构。图 5.3 表示了 4 层树形结构的编码情况。按左支定义为 0，按右支定义为 1。当然，相反的定义也是可以的，因此，前缀码不是唯一的。树中，符号 A 是树根，其他符号按照从根节点出发的路径进行排列。如符号 F 的编码为 10（A→C→F），符号 J 的编码为 010（A→B→E→J）等。实际的编码树一般不是图 5.3 所示的全树，而是全树的子树，从下面的 Huffman 树可以看出这一点。

2. Huffman 编码

Huffman 编码的算法如下：

- (1) 将图像的灰度等级按概率大小进行升序排序。
- (2) 在灰度级集合中取两个最小概率相加，合成一个概率。
- (3) 新合成的概率与其他的概率成员组成新的概率集合。
- (4) 在新的概率集合中，仍然按照步骤 (2)～(3) 的规则，直至新的概率集合中只有一个概率为 1 的成员。这样的归并过程可以用二叉树描述。
- (5) 从根节点按前缀码的编码规则进行二进制编码。

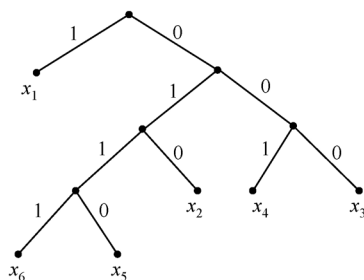
【例 5.1】表 5.2 第一行和第二行列举了一个信源的统计特性，图 5.4 (a) 表示了建立 Huffman 码的过程，图 5.4 (b) 表示从根开始，经各中间节点到叶节点的路径采用二进制编码的情况，结果如表 5.2 的第三行所示。

表 5.2 一个信源的 Huffman 编码

符号集 $\{x_i\}$	x_1	x_2	x_3	x_4	x_5	x_6
概率分布 $\{p_i\}$	0.40	0.20	0.12	0.11	0.09	0.08
Huffman 编码	1	010	000	001	0110	0111

符号集	经排序的 概率分布	第一次 合并后 排 序	第二次 合并后 排 序	第三次 合并后 排 序	第四次 合并后 排 序	第五次 合并后 排 序
x_1	0.40	0.40	0.40	0.40	0.60	1.00
x_2	0.20	0.20	0.23	0.37	0.40	
x_3	0.12	0.17	0.20	0.23		
x_4	0.11	0.12	0.17			
x_5	0.09	0.11				
x_6	0.08					

(a) 编码过程



(b) 采用二进制编码

图 5.4 Huffman 编码示意图

3. Huffman 编码的性能

实现 Huffman 编码的基础是统计源数据集中各信号的概率分布。Huffman 编码在无失真的编码方法中效率优于其他编码方法，是一种最佳变长码，其平均码长接近于熵值。不难计算出上述 Huffman 编码的平均码长是 2.37 比特/字符，接近于信源的熵 2.31。

当信源数据成分复杂时，庞大的信源集致使 Huffman 码表较大，码表生成的计算量增加，编译码速度相应变慢；另外，不等长编码致使硬件译码电路实现困难。上述原因致使 Huffman 编码的实际应用受到限制。为了克服这一缺点，人们研究了亚最佳编码方法，例如在电视编码中可以采用简单的双字长变长编码。这种方法只采用两种字长的码字：对出现频率高的一些符号用短码字（如 3 bit 字长），对出现频率低的一些符号用长码字（如 6 bit 字长），并留一个码字作为长码字的前缀部分。这种编码方式的平均码长尽管不如 Huffman 编码短，但也接近熵值，且硬件的实现起来要简单得多。随着微电子技术的发展，Huffman 编码算法已经被集成到一些编译码集成电路中，并被许多国际标准所采用。

4. 图像的 Huffman 编译码系统

图 5.5 给出了 Huffman 编译码系统的流程。

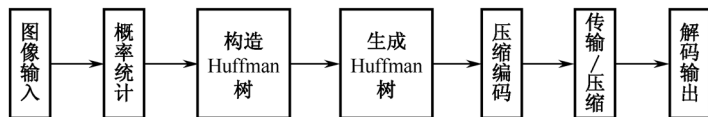


图 5.5 Huffman 编译码系统流程

图中根据 Huffman 算法构造 Huffman 树是 Huffman 编码中的一个难点。由建成的 Huffman 树，便可生成各灰度级对应的 Huffman 代码。方法是依次由根节点出发，前进到每个叶节点，所经路径上对应的“0”或“1”的集合便是各信号对应的 Huffman 码，如图 5.4 (b) 所示。编码压缩模块的功能是将源图像文件转换成 Huffman 代码形式，并存储到另一个文件中或用于信道传输。压缩后的图像文件是一些二进制的“0”和“1”串，且没有分界符。译码模块负责将这些“0”、“1”串参照 Huffman 树恢复为原图像信号。

5.2.2 算术编码

1. 算术编码的基本思想

Huffman 编码采用的思路是一个符号分配一个码字，这种编码方法一般称为分组码。早在 20 世纪 60 年代，R. Elias 就提出了一种与分组码有本质差别的算术编码方法 (arithmetic coding)，但直到 20

世纪 80 年代才得以实现。其基本思想是：按照符号序列的出现概率对概率区间 (interval) 分割，用一个实数代表一个数据流的输入符号，再将这个实数 (二进制小数) 转化为一定位数的二进制代码。对于较长的消息，二进制代码的位数也会增加。它是通过算术运算逐步形成码字的，因而得名。

2. 算术编码的主要步骤

(1) 首先将初始“当前区间”定义为 $[0, 1)$ ，则初始概率空间宽度 A_0 为 1，其下限、上限初始值 C_0 和 D_0 分别为 0 和 1。假设有多个各不相同的符号，已知第 i 个符号出现的概率分配子区间为 $[p_{ih}, p_{il})$ ，概率为 $p_{ih} - p_{il}$ 。

(2) 对输入流 $s_1 s_2 s_3 \dots$ 中的每个符号 s_n ($n = 1, 2, 3, \dots$)，重复下面的两步：

① 把当前区间分割为长度正比于符号概率的子区间。设上次迭代的概率空间宽度为 A_{n-1} ($n = 1, 2, 3, \dots$)，本次迭代过程的概率空间的下限值和上限值分别为 C_n 和 D_n ，迭代公式为

$$C_n = C_{n-1} + A_{n-1} \cdot p_{nl} \quad (5.12a)$$

$$D_n = C_{n-1} + A_{n-1} \cdot p_{nh} \quad (5.12b)$$

② 为 s_n 选择一个子区间，并将其定义为新的当前区间。

(3) 当把整个输入流处理完后，输出的即为能唯一确定当前区间的数字。

【例 5.2】 输入的符号序列为 0110。通过简单乘法和加法的递归运算，求 n 个符号输入后的概率区间宽度 A_n 、下限值 C_n 以及上限值 D_n 。

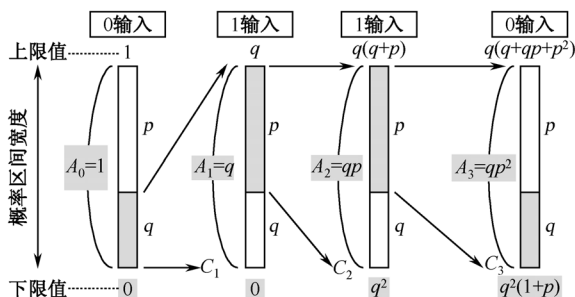


图 5.6 符号序列 0110 的算术编码过程

(最后一个 0 输入下面原来的 $q(p+pq+p^2)$ 改为 $q^2(1+p+p^2)$)

如图 5.6 所示，符号 1 的发生概率为 p ，符号 0 的发生概率为 q ($q < p$, $q + p = 1$)，则概率子空间范围的划分为：符号 0 为 $[0, q)$ ，符号 1 为 $[q, 1)$ 或 $[1-p, 1)$ 。首先，以发生概率 q 和 p 的比率对概率空间宽度 A_0 进行内分，将对应于已被输入符号值的概率空间宽度设置为新的概率空间宽度，与此同时，重新设置概率区间下限值。

由于 $s_1 = 0$ 最先被输入，所以，下一个概率空间宽度 A_1 为 q ，其下限值 C_1 为 0，没有变

化，按照实际输入值逐次将这样的程序执行下去。迭代过程如下：

设定初始值： $A_0 = 1$, $C_0 = 0$

输入 $s_1 = 0$ 后： $A_1 = A_0 q = q$, $C_1 = C_0 = 0$, $D_1 = C_0 + A_1 = q$

输入 $s_2 = 1$ 后： $A_2 = A_1 p = qp$, $C_2 = C_1 + A_1 q = q^2$, $D_2 = C_1 + A_2 = qp$

输入 $s_3 = 1$ 后： $A_3 = A_2 p = qp^2$, $C_3 = C_2 + A_2 q = q^2(1+p)$, $D_3 = C_2 + A_3 = q$

输入 $s_4 = 0$ 后： $A_4 = A_3 q = q^2 p^2$, $C_4 = C_3 = q^2(1+p)$, $D_4 = C_3 + A_4 = q^2(1+p+p^2)$

这样得到最后的一组概率范围 $[C_4, D_4]$ ，编码结果结束。假设 $p = 2/3$, $q = 1/3$ ，则最后的范围是 $[5/27, 19/81) = [0.185185, 0.234568)$ 。

解码过程如下：

解码器得到最后的范围，根据各符号的概率值及其所分配的范围，只有符号 0 产生的范围能包含 $[0.185185, 0.234568)$ 。解码器范围的初始值为 $[0, 1]$ ，在解码出 0 后，范围变为 $[0, 1/3]$ 。即在公式 (5.12) 中范围进行了更新， $A_{n-1} = 1/3$ ，重新计算两个字符的产生范围，得出概率子空间范围的划分为：符号 0 为 $[0, 1/9)$ ，符号 1 为 $[1/9, 1/3)$ ，并与最终范围进行比较，解出第二个符号为 1。余可类推，解码器可以唯一解出输入流 0110。

算术编码的特点:

(1) 该方法实现起来不如 Huffman 方法简单, 因此应根据具体的图像类型来选择相应的编码方法。算术编码过程不采用像 Huffman 编码那样给符号值分配整数码字的方法, 而是把二进制数所代表的概率空间宽度叠加到代码串中去, 并由这样的处理形成算术编码过程。

(2) 虽然算术编码的硬件实现比 Huffman 编码方法复杂, 但对图像的编码测试结果表明, 在信源符号概率接近的条件下, 算术编码效率能提高 5% 左右的效率, 明显高于 Huffman 方法。因此, 在扩展的 JPEG 系统中用算术编码取代了 Huffman 编码。

(3) 算术编码除了常见的基于概率统计的模式外, 还有自适应模式。在这种模式下, 各符号的初始概率相同, 它们依据出现的符号而发生变化。这种模式特别适用于不便于进行符号概率统计的实际场合。实际应用场合, 常常不可能长时间对全部大量的数据进行概率同统计, 而短时间统计得出的固定概率模型对不同的输入流的编码效率不一定很高, 算术编码的自适应模式正好弥补了这一缺陷。

5.3 预测编码

预测编码(predictive coding)应用了现代统计学和控制理论的“时间序列分析”概念, 是有损压缩中的重要方法之一。早在 20 世纪 40 年代, Wiener 等人就对线性预测开展了开创性的工作。现在多媒体计算机系统语音分析与合成、图像的编码与解码等都广泛地采用了预测编码。

预测编码的基本思想是: 在某种模型的指导下, 根据过去的样本序列推测当前的信号样本值, 然后用实际值与预测值之间的误差值进行编码。如果模型与实际情况符合得比较好且信号序列的相关性较强, 则误差信号的幅度将远远小于样本信号。对电视图像的统计表明, 原图像幅度的概率分布大体比较均匀, 而经预测编码后的差值绝大部分处在零附近的极小范围内, 只有在图像的轮廓和边缘处出现较大的预测误差。这样, 就可以用较少的量化层次和比特数表示小幅度的误差信号, 压缩了数据; 至于偶尔出现的大幅度误差信号, 因为人眼对跳变的边沿误差并不敏感, 也可以减少量化级数。

5.3.1 预测编码基本原理

预测编码是利用图像的空间或时间的冗余度进行的。对实际值与预测值之间的误差值进行编码的方法称为差分脉冲编码调制(Differential Pulse Code Modulation, DPCM)。图 5.7 表示了 DPCM 系统组成的原理, 图中假设信道无误码地传输或存储数据。

图中 x_n 表示 t_n 时刻的信号取样值。预测器根据 t_n 前的样值 $x_1, x_2, x_3, \dots, x_{n-1}$ 对 t_n 进行预测, 得出 x'_n 与 x_n 之间的误差值 e_n

$$e_n = x_n - x'_n \quad (5.13)$$

对 e_n 进行量化得出 e'_n 。编码器对 e'_n 进行编码发送。经译码后, 接收端恢复出信号 x''_n 。恢复误差为

$$x_n - x''_n = x_n - (x'_n + e'_n) = e_n - e'_n \quad (5.14)$$

当上述差值为 0 时, DPCM 系统可以做到无失真地恢复原始信号序列, 但事实上, 预测、量化等误差总是存在的。根据预测方案的不同, 预测编码可以分为线性预测编码、非线性预测编码等。

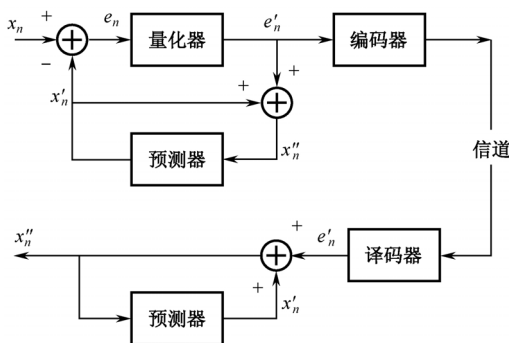


图 5.7 DPCM 系统的组成

5.3.2 线性预测编码

在 DPCM 中, 预测器十分重要, 其预测的准确程度直接影响着压缩性能。采用线性预测函数在理论上较为成熟且便于实际应用。假设经扫描后的图像信号 $x(t)$ 是一个均值为零、方差为 σ^2 的平稳随机过程。线性预测就是选择 a_i ($i = 1, 2, \dots, N-1$) 使预测值为

$$x'_n = \sum_{i=1}^{N-1} a_i x_i \quad (5.15)$$

并且使差值 e_n 的均方值为最小。

预测信号的均方误差 (MSE) 定义为

$$E\{e_n^2\} = E\{(x_n - x'_n)^2\} \quad (5.16)$$

为了设计最佳预测的系数 a_i , 采用最小均方误差准则 (MMSE)。可以令

$$\frac{\partial E\{e_n^2\}}{\partial a_i} = 0 \quad (5.17)$$

定义 x_i 和 x_j 的自相关函数

$$R(i, j) = E\{x_i x_j\} \quad (5.18)$$

在序列为平稳随机过程的条件下, $R(i, j) = R(i-j)$, 则

$$R(i) = \sum_{k=1}^{N-1} a_k R(k-i) \quad (5.19)$$

式中 $i = 1, 2, \dots, N-1$ 。可将式 (5.19) 写成矩阵形式为 Yule-Walker 方程组

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(N-2) \\ R(1) & R(0) & \cdots & R(N-3) \\ \vdots & \vdots & \cdots & \vdots \\ R(N-2) & R(N-3) & \cdots & R(0) \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{N-1} \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(N-1) \end{bmatrix} \quad (5.20)$$

在 $R(i)$ 已知的条件下, 该方程组可以用递推算法来求解 a_i 。

通过分析可以得出以下结论:

(1) 压缩效果与图像的相关性密切相关。图像的相关性越强, 压缩效果越好。

(2) 预测阶数并非越大越好。当某个阶数已使 $E\{e_N e_{N-1}\} = 0$ 时, 即使再增加预测点数, 压缩效果也不可能继续提高。

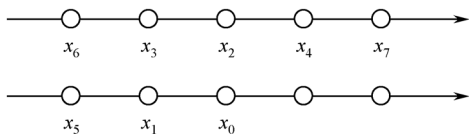


图 5.8 当前像素与邻近像素的位置关系

(3) 若 $\{x_i\}$ 是平稳 m 阶马尔可夫过程序列, 则 m 阶线性预测器就是在 MMSE 意义下的最佳预测器。

在前面的讨论中, 对 x_1, x_2, \dots, x_{N-1} 的取法没有做具体规定。实际上, 常用的预测器有多种方案。图 5.8 表示了当前像素 x_0 与邻近像素 x_i 之间的位置关系。

(1) 前值预测: 用 x_0 同一行的最近邻近像素来预测, 即

$$\hat{x} = x_0 \quad (5.21)$$

实际进行预测时, 如果用公式 (5.15) 估算当前值, 则要求用到前 $N-1$ 个样值, 且依赖于输入端的统计特性。可见, a_i 的获取较难, \hat{x} 的计算量也较大。采用前值预测简单方便, 编码效率高, 当然压缩性能稍逊。JPEG 标准就采用前值预测来实现压缩。

(2) 一维预测: 用 x_0 同一行的前若干邻近像素来预测, 如图 5.4 中的 x_1 和 x_5 。

(3) 二维预测: 用 x_0 同一行的前面所有邻近像素和前几行的取样值来预测, 如图 5.8 中的 x_1 、 x_2 、 x_3 、 x_4 、 x_5 、 x_6 、 x_7 等。

(4) 三维预测: 在二维预测的基础上, 利用上帧或前几帧的邻近取样值作为 x_0 的取样值。这种预测方法主要用于视频图像的压缩, 它被体现在 MPEG 标准中。

5.3.3 自适应预测编码

上面介绍的 DPCM 系统, 当输入信号为平稳的样本序列时, 效果较好。但当输入数据总体平稳而局部有较大变化时, 使用固定参数的预测器将会影响预测的准确性。为了适应实际信号变化的需要, 可以根据信号的变化规律自适应地改变量化和预测特性。此时的预测编码方案称为自适应差分脉冲编码调制 (ADPCM) 预测编码。

自适应量化编码指系统能随信号样值的改变而自动改变量化区间的大小以保持输入量化器的输入信号基本均匀。ADPCM 必须有确定输入信号变化量的估值能力。为了达到实时化的要求, 信号的估值必须实现简单。若估值在信号的输入端进行, 称为前馈自适应; 若估值在量化器的输出端进行, 称为反馈自适应。

自适应预测是指预测参数根据信号的统计特性来确定, 以达到最佳预测。前面介绍的 DPCM 采用固定的预测参数往往预测效果较差, 要得到最佳预测参数计算量较大。实际的自适应预测器是将两种方法折中考虑, 预测参数仍采用固定数值, 但根据常见的信源特性提供可选的多组预测参数。例如一种 ADPCM 方案是: 采用二样值预测, 提供 7 组预测系数, 编码时根据选定的 MMSE (最小均方估计) 准则, 每个编码区间自动选取一组最佳的参数以达到最佳预测。

为了进一步提高压缩比, 非线性自适应预测充分考虑了图像的统计特性及其变化, 尽量使预测系数与图像所处位置的局部特性相匹配。最佳非线性预测的思想是: 当 x_0 的邻近像素的联合概率密度 $p(x_0 | (x_1, x_2, \dots, x_n))$ 已知时, 可以求得 x_0 的最佳预测值 \hat{x}_0 。可见, 非线性预测是一种自适应预测, 它反映了相邻像素之间的制约关系, 使图像的局部特性得到良好的匹配。

预测编码的优点是直观快捷、便于实现, 特别适用于具有实时性的硬件结构, 在传输速率较高的场合大都采用这种方法。预测编码的缺点是压缩比不够高。

5.4 变换编码

变换编码是另一种经典的数据压缩的基本方法。正交变换使图像的表达形式发生了变化, 同时为图像的高压缩比提供了可能性。本节将介绍变换编码的思想方法。

5.4.1 变换编码的基本原理

通过对信号进行某种函数变换, 实现从信号相关性较强的信号空间 (如一维时间域、二维空间域) 变换到像素相关性较弱、便于编码的另一个信号空间 (如频率域)。它的关键是选择与信号空间特征相匹配的变换函数。通过数学变换可以改变信号能量的分布, 从而压缩信息量。这些数学变换可以是简单的坐标转换, 也可以是时间域、空间域的正交变换。

以傅里叶变换的概念说明变换可以改变信号能量的分布的基本原理。一个时间域的正弦波的电压瞬时值各不相同, 而把它转换到频率域后信息相对集中, 表示它就相当简洁, 只有一根谱线。相反, 时间域上的脉冲波形, 看上去并不复杂, 但当把它转换到频率域上时, 新的表示形式反而变得不直观。

所以，如果选定的一个变换域与图像的特征相匹配，就可以简化编码，大大压缩码率。这里的频率反映了周期型图案的主要特征，而脉冲波形是图像中出现边沿的抽象，用 Haar 小波基来表示会更加适宜。

5.4.2 变换编码的系统结构

与预测编码相比较，预测编码在时空域进行，而变换编码主要在变换域进行。典型的变换编码系统组成如图 5.9 所示。

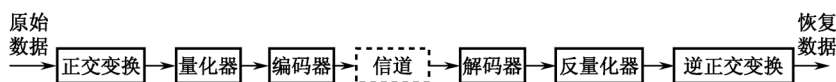


图 5.9 变换编码系统原理框图

图像数据经过线性正交变换，从原先彼此密切相关的像素（空间域）变换为统计上彼此较为相互独立、甚至达到完全独立的变换域系数矩阵。由于正交变换是可逆的，变换前后的信息熵保持不变，这样去相关后，在变换域中的一阶熵可以近似图像的高阶条件熵，用变长字编码就能达到较好的压缩。

一般信号的能量主要集中在低频部分。在上述系统中，如果对能量较小的高频部分不编码或仅分配较少的比特数，就可以明显地减少传输或存储的数据量。对变换系数的编码也可以采用非线性量化器，以提高编码效率。

5.4.3 变换编码方案的选取

在所有的变换编码方案中，离散 K-L 变换（Hotelling 变换）是最佳变换，理论价值较高，常常作为对其他变换特性进行评价的标准。但此变换没有快速算法，在工程应用中受到限制。

在次最佳变换算法中，DFT（离散傅里叶变换）和 DCT（离散余弦变换）都是常用的变换编码方法，它们分别有快速算法：FFT（快速傅里叶变换）和 FCT（快速余弦变换）。这两种方法相比较，DFT 涉及复数运算，而 DCT 是实数变换，具有十分吸引人的一些特点：它是一种实数变换，计算量较小，其变换矩阵的基向量很好地描述了人类视觉的相关性，且对于大多数图像来说，该变换的压缩性能很接近离散 K-L 变换，而且其变换矩阵与图像内容无关。另外，由于它构造对称的数据序列，避免了在图像边界处的跳跃及所引起的吉布斯效应，并且也有快速算法，因而得到广泛的应用。作为准最佳变换，它已成为一些静态图像、视频压缩国际标准（或建议）中的基本处理模块。JPEG 制定了基于 DCT 的有失真静止图像压缩标准。在 MPEG 视频编码器中，帧内图像（I 图像）采用了 DCT 的编码方法。

在变换编码中，如下几个问题值得注意。

1. 子图像大小的选取

利用正交变换进行压缩编码时，需要考虑的一个实际问题是实现的复杂性。一般的图像尺寸都比较大，进行全尺寸的直接变换计算量太大，尤其对于硬件实现能以承受，因此综合实现的复杂性和编码效率等因素，在包含现有的国际标准在内的实际应用中，均采用 $n \times n$ 子图像变换。我们把这种变换编码称为基于子块的正交变换编码。将大小为 $M \times N$ 像素的图像分解成 MN/n^2 个 $n \times n$ 子图像，再对子图像进行变换。 n 取得越小，越便于实时处理。但子图像越小，压缩效果越差，且在图像复原时图像的整体感越差，易出现子图像之间不能较好衔接的“方块效应”或“马赛克效应”。

图像一般在相邻的 20 个像素之间存在着相关性，所以若 $n > 16$ 时再增加 n 的值，则对改善图像质量作用不明显，相反会增加变换的计算量，且硬件的复杂程度与子图像的大小成正比。一般选择 $n = 8$ 或 16。

JPEG 编码标准先将图像分成 8×8 大小的子图像，然后对各个子块分别进行处理，包括 DCT 系数量化、直流系数的预测、交流系数的 Z 字形扫描和行程编码等，以取得更高的压缩效率。

2. 变换方法的选取

有三种常用的正交变换编码。

(1) 区域编码

变换域的能量主要集中在低频区域，而大部分区域的能量为零或可以忽略不计。选取能量集中的区域进行编码，舍弃能量为零或很小的区域，从而达到数据压缩的目的。

区域编码适用于原始图像经过滤波且滤波器频率特性已知的情况。该编码方法压缩比可达 5:1，但由于高频分量被丢弃，致使图像的分辨率下降。由于变换域的能量分布不十分清楚，上述保留区域和丢弃区域的划分有时难以判断。

(2) 阈值编码

规定一个阈值，对大于阈值的变换系数幅度进行编码，舍弃在阈值以下的变换系数分量。这样不仅保留了能量较大的低频分量，而且某些有一定幅度的高频分量也能得到保留。所以，这种编码方法有效地改善了重建图像的质量。

阈值编码要对变换系数在二维频谱中的位置进行编码，因而要比区域编码复杂。如果能根据子图像的细节和灰度分布情况自动调整阈值，则可实现自适应阈值编码。

将变换编码与其他编码方法一起进行混合编码是十分有意义的。如 JPEG 标准中允许 DCT 变换编码与预测编码、熵编码以及下面将要讲到的行程编码混合使用，可以在满足图像质量的前提下达到很高的压缩比。

【例 5.3】对给定图像进行二维 DCT 变换，并将 DCT 系数矩阵中值小于给定阈值 ($T_1 = 5$, $T_2 = 50$) 的元素置为 0，然后使用反离散余弦变换重建图像，并计算各自的均方误差。

【解】程序如下：

```
F = imread('flowers.tif');
figure(1); imshow(F);           %显示原始图像
G = rgb2gray(F);                %将 RGB 图像转换为灰度图像
dctG = dct2(G);
figure(2); imshow(log(abs(dctG)), []); %显示 DCT 系数图像
T1 = 5; T2 = 50;
dctG(abs(dctG)<T1) = 0;
idctF1 = idct2(dctG);
figure(3); imshow(idctF1,[0 255]); %显示重建图像
dctG(abs(dctG)<T2)=0;
idctF2=idct2(dctG);
figure(4); imshow(idctF2,[0 255]); %显示重建图像
[M N] = size(G);
MSE1 = sum(sum((idctF1-double(G)).^2))/(M*N) %重建图像 1 的均方误差
MSE2 = sum(sum((idctF2-double(G)).^2))/(M*N) %重建图像 2 的均方误差
```

运行结果如图 10 所示。

MSE1 = 3.3058

MSE2 = 152.8322

可见，低阈值压缩 ($T_1 = 5$) 图像高频分量损失较小，主观难以观察到画质的变化；而较高的阈值 ($T_2 = 50$) 压缩后复原的图像已出现明显的失真。

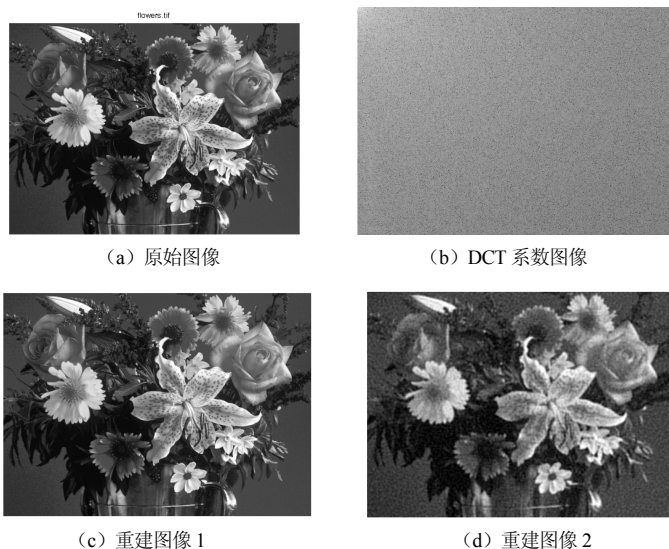


图 5.10 利用 DCT 进行图像压缩

5.4.4 整数小波变换与图像压缩

由于量化器的设计是决定图像保真度的关键环节，而传统的 DCT 和第 4 章讲述的经典小波变换（又称为第一代小波）在图像变换后会产生浮点数，因而必须对变换后的数据进行量化处理，这样就产生了不同程度的失真。新一代的整数小波变换（又称为第二代小波变换）采用提升方法能够实现整数变换，因而能够实现图像的无损压缩，显然它是一种很适合于医学等图像的压缩方法。

提升方法构造小波分为分裂、预测和更新 3 个步骤。下面以一维数据为例加以说明。

1. 分裂 (split)

将一原始信号序列 s_j 按偶数和奇数序号分成两个较小、互不相交的小波子集 s_{j-1} 和 d_{j-1}

$$\text{split}(s_j) = (\text{even}_{j-1}, \text{odd}_{j-1}) = (s_{j-1}, d_{j-1}) \quad (5.22)$$

分裂步骤结束时，它们等于原始信号的偶数位和奇数位的数列。

2. 预测 (predict)

由于数据间存在相关性，因而可以定义一个预测算子 P ，用 $P(s_{j-1})$ 来预测 d_{j-1} ，这样可用相邻的偶数序列来预测奇数序列。若用 d_{j-1} 与 $P(s_{j-1})$ 的差值代替 d_{j-1} ，则其数据量要比原始 d_{j-1} 小得多

$$d_{j-1} \leftarrow d_{j-1} - P(s_{j-1}) \quad (5.23)$$

其中，“ \leftarrow ”表示赋值运算。

一种简单的情况如下：取两个相邻偶数序号所在数据的均值作为它们间奇数序号所在数据的预测值，即

$$P(s_{j-1}) = (s_{i,2k} + s_{i,2k+2}) / 2 \quad (5.24)$$

预测步骤结束时， s_{j-1} 不变， d_{j-1} 等于原始 d_{j-1} 信号与预测值的差。

3. 更新 (update)

由于上述两个过程一般不能保持原图像中的某些整体性质（如亮度），为此我们要构造一个 U 算子去更新 s_{j-1} ，使之保持原有数据集的某些特性

$$s^{j-1} \leftarrow s^{j-1} + U(d_{j-1}) \quad (5.25)$$

若取

$$U(d_{j-1}) = (d_{j-1,k+1} + d_{j-1,k}) / 4 \quad (5.26)$$

由此推算出的提升方法称为 (5,3) 小波变换, 变换公式为

$$\begin{cases} d_{j-1,k}^- = (s_{j-1,k} + s_{j-1,k+1}) / 2 \\ s_{j-1,k}^+ = (d_{j-1,k+1} + d_{j-1,k}) / 4 \end{cases} \quad (5.27)$$

逆变换为

$$\begin{cases} s_{j,2k} = s_{j-1,k} - (d_{j-1,k-1} + d_{j-1,k}) / 4 \\ d_{j,k+1} = d_{j-1,k} + (s_{j,2k} + d_{j,2k+2}) / 2 \end{cases} \quad (5.28)$$

实现图像的整数小波变换和逆变换, 需要对行和列各做一次。采用最简单的 S 小波, 它是 Haar 变换的整数形式, 数学表达如下。

正变换

$$d_{n+1,l} = s_{n,2l+1} - s_{n,2l} \quad (5.29)$$

$$s_{n+1,l} = s_{n,2l} + \left\lfloor \frac{d_{n+1,l}}{2} \right\rfloor \quad (5.30)$$

其中, s 代表变换前后的图像数据, n 和 l 代表行和列, d 为中间变量, 符号 $\lfloor \cdot \rfloor$ 为取整算符。

逆变换为

$$s_{n,2l} = s_{n+1,l} - \left\lfloor \frac{d_{n+1,l}}{2} \right\rfloor \quad (5.31)$$

$$s_{n,2l+1} = d_{n+1,l} + s_{n,2l} \quad (5.32)$$

正变换程序如下, 结果分别如图 5.11 至图 5.14 所示。



图 5.11 奇偶列重排的结果

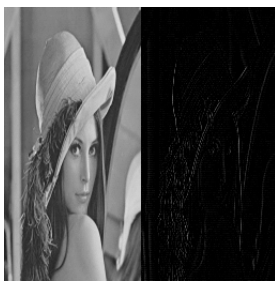


图 5.12 行变换的结果

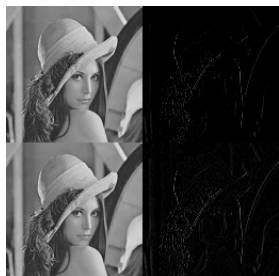


图 5.13 奇偶行重排的结果

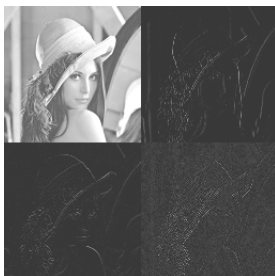


图 5.14 列变换的结果

5.5 二值图像编码

只用“白”（用“0”表示）和“黑”（用“1”表示）两个灰度级表示的图像称之为二值图像（binary image）。两个色彩的图像也可以用二值图像表示。自然界的图像一般不是二值图像，通常它们是人为产生的，如由文字组成的文档文件、表格、工程图纸、地图等。由于二值图像具有两个灰度级，因此可以充分利用其统计特性采用相对简单的编码压缩方法。一幅二维图像按位平面进行分解可以得到若干二值图像，因而二值图像的编码方法为促进编码模式奠定了基础。

5.5.1 跳跃空白编码

对二值图像进行编码最直接的方法是用一位二进制码描述一个像素。显然，这时编码文件的比特数就是图像的像素数。随着二值图像分辨率的提高，相应的数据量也成比例地提高。

跳跃空白编码（skip blank coding）将图像的每一条扫描线分成若干等长的段，每段有 m 个像素，一般 $m = 8 \sim 12$ 。这些扫描线段的组成可能出现两类情况：

（1）全是“0”像素。这种线段称为“空白块”（blank），如两行文字中间的空白部分。编码时“空白块”用码字“0”表示。在二值图像中，“空白块”常常表示二值图像的背景成分，总要占有相当大的比例。

（2）全是“1”像素或由“0”、“1”像素混合而成。编码时，这种线段用“1”加直接编码表示。

尽管上述编码中非“空白块”线段在首位要附加一个“1”，但可以证明只要空白块出现的概率大于 $1/m$ ，跳跃空白编码就能达到压缩数据量的目的。

上述方法很容易推广的在二维情形中。将图像划分为若干子图像。当子图像像素全为0时，编码为“0”；否则，子图像的编码首位为“1”，其余码位（code position）采用像素的直接编码表示。

5.5.2 游程长度编码

游程长度编码（Run Length Coding, RLC）或称行程编码是一种简单的无损编码技术，它改变连续出现相同字符的表达方式，以降低码长。如前所述，在传真的二值图像中，连“0”或连“1”总是成串出现，称为“白游程”和“黑游程”，例如5个连“0”可以传送“白5”。其他非二值的相同连续数据串，同样简化为两个符号：一个符号代表数据，另一个符号代表串长。

游程长度编码一般不直接单独使用，通常配合其他编码方式使用来提高压缩效果。

除全白或全黑扫描行外，在二值图像扫描行上，黑、白像素段总是交替产生的。对由信号 x_1, x_2, \dots, x_N （ N 为扫描行中像素的最大序号）组成的二值图像扫描行，可以视为若干白像素段（白长）和黑像素段（黑长）之和。游程编码根据二值图像扫描行的这些特点，采用将黑长和白长以游程（RL）分别分配码字的分配方法。根据交替的性质，第一个线段的性质指明以后，后面线段的性质就可以判断出来。一种编码方法是：第1段线段游程长度，第2段线段游程长度，第3段线段游程长度……。

游程编码需要计算图像信号出现的行程长度，再将行程长度转换为代码。理论上，应该根据霍夫曼算法按RL出现的概率来分配码字。但实际上，不同图像中的同一个RL出现的概率并不一致。为了用传真机实时传送各种二值图像，避免概率统计花费较多的时间，国际电报电话咨询委员会（CCITT）第15研究组制定了T.4建议，该建议推荐三类文件传真机以8幅典型传真文件的标准样张为统计依据编出霍夫曼码表，称为改进的霍夫曼编码（Modified Huffman Coding, MHC）。

CCITT 传真压缩标准为保证传真文件有足够的清晰度，规定A4（210 mm × 297 mm）应该有1188或2376条扫描线，相当于4线/mm或8线/mm。每条扫描线取样点为1728点，相当于8 pel/mm或200 dpi。G3、G4中的一维压缩方法采用了变长行程编码方法。编码规则如下：

- (1) 全组码字由结束代码、组成代码两部分组成。
- (2) $RL = 0 \sim 63$, 行程用相应的结束代码字表示。根据统计, 此类 RL 占多。
- (3) $RL = 64 \sim 1728$, 首先由最接近实际行程长度的组成代码编码, 后再跟结束代码字。
- (4) 每行从白游程开始。若实际扫描行由黑游程开始, 则需在行首加 $RL = 0$ 的白游程。
- (5) 每行结束要加行同步码 EOL。

CCITT 也对行像素宽展到 2623 的加宽纸型的二值图像做了补充规定:

- (1) 全组码字由结束代码、组成代码和附加组成码三部分组成。

- (2) $RL = 0 \sim 63$, 行程用相应的结束代码字表示。

(3) $RL = 64 \sim 2623 \times (2560 + 63)$ 时, 首先由最接近实际行程长度的组成代码编码, 后再跟结束代码字。

(4) $RL \geq 2624$, 首先按上述方法由组成代码 2560 编码, 如果行程的剩余部分大于 2560, 则按规定的附加组成代码, 直到行程的剩余部分小于 2560, 再按前述办法编码。

上述编码方法是在每个扫描行内进行的, 称为一维 MH 编码。相应地, 如果利用二值图像在主、副两个扫描方向上的相关性, 还可以进一步提高二值图像的压缩比。

5.6 图像压缩编码标准

在静态图像压缩编码标准中, 比较著名的有 JPEG、JBIG 等标准。事实上, 视频可以视为一幅幅不同但相关的静态图像的时间序列。因此, 静态图像的压缩技术和标准可以直接应用于视频的单帧图像。本节首先介绍适用于静态图像的 JPEG 标准和 JBIG 标准, 然后介绍新的 JPEG2000 压缩国际标准。

5.6.1 彩色与灰度图像压缩标准 JPEG

JPEG 标准是由国际标准化组织 (ISO) 和 CCITT 的联合图片专家组 (Joint Picture Expert Group, JPEG) 针对静态数字图像压缩而联合制定的第一个国际标准, 适用于压缩各种分辨率与格式的灰度图像和彩色图像。和相同图像质量的其他常用文件格式 (如 GIF、TIFF、PCX) 相比, JPEG 格式对静态图像的压缩要高得多, 所以在数码相片、彩色照片的网络传输等方面应用非常广泛, 成为最流行的图像格式。正是由于其高压缩比, 使得 JPEG 被广泛地应用于多媒体和网络程序中, 例如 HTML 语法中选用的图像格式之一就是 JPEG。尽管现在已有更新、压缩比更高 (压缩比比 JPEG 高约 30%)、功能更强的 JPEG2000 算法与标准, JPEG 图像仍在彩色图像压缩方面占主要地位。

1. JPEG 算法与系统

(1) JPEG 基本系统: 采用 DCT 算法的顺序模式实现图像的有损压缩。整个编码采用常规顺序扫描方式一次完成。该系统的性能依赖于图像的复杂性, 对一般图像常取 $20:1$ 或 $25:1$ 。在压缩比为 $25:1$ 的情况下, 解压缩后还原的图像与原图像相比较, 人眼视觉难以区分它们的差别。该系统应用十分广泛, 以至于人们认为 JPEG 标准一般即指其基本系统。标准 JPEG 又被称为基线 (baseline) JPEG, 在网页中需要全部下载后才能观看整个图像。

(2) JPEG 扩展系统: 包含了基本系统的算法, 增加了类似于精细可伸缩性 (FGS, Fine Granular Scalable) 编码技术的“逐渐浮现”的累进 (processive) 方式“由粗而细”地传送或显示图像, 因而图像每个分量的编码要经过多次扫描才能完成。第一次扫描只传送较粗糙的压缩编码, 然后解码显示该质量较低的图像。如果该图像满足要求, 则停止扫描; 否则在进行第二次扫描时做较细的压缩。不断继续上述过程, 直到满意为止。可见, 扩展系统为通信用户提供了一种灵活选择图像质量和传输时间的方式。

当网速较慢浏览网页时,可以观察到基线格式的图片一行行的出现且很清晰,而累进格式的图片开始出现一大块有很多看不清楚马赛克图案,然后逐渐清晰起来。

(3) 信息保持压缩系统:即无损压缩模式。利用图像在空间域存在的相关性,采用线性预测并对预测误差进行霍夫曼编码或算术编码实现图像的无损压缩,其图像质量高,但压缩比要比基本系统采用的有损压缩低得多,对于中等复杂度的彩色图像,压缩比约为 2:1。

2. JPEG 基本系统

基本系统的定义保证了压缩比和压缩质量。每个单独的彩色图像分量的编码算法如下。

(1) 将量化精度为 8 位的待压缩图像分成若干 8×8 的样值子块,做基于 8×8 子块的 DCT。

(2) 根据最佳视觉特性构造量化表,设计自适应量化器并对 DCT 的频率系数进行量化。

(3) 为了增加连续的“0”系数的个数,对量化后的系数进行 Z 字形重排。对直流系数进行 DPCM 编码,对交流系数进行 RLC 编码。

(4) 使用霍夫曼码作为变长熵编码器,对量化系数进行编码,进一步压缩数据量。JPEG 提供了直流系数、交流系数码表,对亮度分量和色度分量采用不同的码表。

解码过程与编码过程正好相反。上述算法的几点说明如下。

(1) 彩色空间转换问题:JPEG 算法处理的是单独的彩色分量图像,所以来自其他彩色空间的图像数据要以 JPEG 格式保存,需要进行彩色空间的转换(见第 10 章),如将 RGB 空间、 $Y_C R_C B_C$ 空间或转换为 YUV 空间等。

(2) 量化:量化的目的是为了压缩数据,同时也是图像质量下降的主要原因。所以设计合理的量化器十分重要。在保证图像质量的前提下,为了获得较高的压缩比,JPEG 量化器利用人眼的空间视觉特性,相对于高频成分对低频成分采用较小的量化间隔和较少的比特数。又根据人眼对亮度信号比色度信号敏感的原理,对图像的亮度分量和图像的色差分量使用不同的量化表——亮度量化表和色差量化表。量化表的元素即为量化间隔。对于 CCIR 601 标准电视图像,JPEG 标准提供了最佳的亮度和色度量表。根据不同的应用需要,用户还可以设计或选择其他的量化表。

JPEG 的量化采用均匀量化器。量化公式为

$$C_Q(u,v) = \text{Integer} \left\{ \text{Round} \left[\frac{C(u,v)}{Q(u,v)} \right] \right\} \quad (6.33)$$

式中, $C(u,v)$ 是 DCT 系数, $Q(u,v)$ 是量化间隔, Round 表示舍入运算, Integer 表示取整运算。

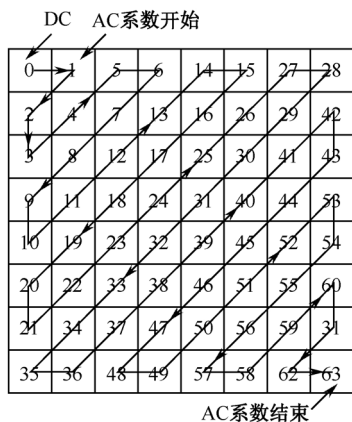


图 5.15 DCT 系数的 Z 字形排列

(3) Z 字形重排: DCT 系数左上角(第 1 行第 1 列)为直流分量(DC 系数),对 8×8 子块矩阵进行 Z 字形编排则可将其余的交流分量(AC 系数)按“频率”从低到高排列,形成 1×64 的向量,如图 5.15 所示。这样排列可以增加“0”系数的游程长度,提高压缩效率。

(4) DCT 系数的编码:由于直流系数表示子块的平均亮度,所以相对于子块图像的交流系数,直流系数较大且相邻块之间的差异较小。JPEG 选择差分脉冲调制(DPCM)编码方式。对相邻块之间的 DC 系数的差值进行编码以获得良好的压缩效果。经过 Z 字形排列后,AC 系数含有较多的“0”系数,因而采用游程长度编码对交流系数进行编码十分有效。

(5) JPEG 位数据流:为了便于 JPEG 格式图像的存储、传输和显示,JPEG 编码器最后需将各种标记码(tag)和压缩数据按帧

组成位数据流。标记码类似于其他图像格式的头信息，给出了 JPEG 图像的所有信息，如图像的宽、高、霍夫曼表、量化表等。标记码有很多，但绝大多数的 JPEG 文件只包含几种。

3. JPEG 图像的画质

标准 JPEG 是一种很灵活的图像存储格式，具有调节图像质量的功能，允许用不同的压缩比（如 10~40）对图像进行压缩。当然，我们要在图像质量和文件大小之间折中考虑。

从保真度的角度看，有些单反相机中 JPEG 记录画质和 RAW 记录画质有很大区别，但在现有的显示器上难以用肉眼分辨。JPEG 是一种有损压缩算法，如果色温不对时，JPEG 格式很难调到准确。如果拍的是重要的、不容失真的场景，最好是用 RAW 数据存储。RAW 是原始的无损数据，并非图像格式，数据量很大且需专门的软件才能显示。RAW 数据记录的层次、颜色、细节可达到 14~16 位，有很好的后期调整余地。

有些数码相机里有 JPEG 画质调节，用不同的圆弧表示不同的存储质量：圆滑的半圆图标表现精细画质，轻微锯齿的半圆图标表现标准画质和明显锯齿的图标表示基本画质，这些图标生动地表现了图片信息损失程度。这里画质的调节本质上是选择 JPEG 压缩率的高低，如精细存储的压缩比约为 4，标准存储约为 8，基本存储约为 16。因为 JPEG 是压缩格式，每次编辑都会进行压缩，所以多次编辑后画质下降就很严重，而 RAW 是非压缩的，保留的细节更多，更适合于多次编辑的照片。

如图 5.16 所示，Photoshop 提供的“JPEG 选项”中的“图像选项”可提供用户控制图像质量（品质 Q），JPEG 格式以使用户选择不同细节的储存需求，以 0~10 级表示，并分成最佳（Q=10~12）、高（Q=8~9）、中（Q=5~7）和低（Q=0~4）4 个档次。其中，0 级压缩比最高，但图像品质最差。一组典型的测试数据可见一斑：Q=12、8、5 和 0，相应的数据量为 309.5 kB、115.8 kB、88.8 kB 和 56.1 kB。采用几乎无损的 10 级质量保存时，压缩比也可达 5。一般，采用第 8 级压缩为存储空间与图像质量兼得的最佳比例。



图 5.16 PS 中 JPEG 格式储存选项界面

JPEG 是网页设计常用的格式。图 5.16 还针对网页提供了不同的“格式选项”。对于网页来说，图像大小至关重要，显示速度是第一位的。按“基线（“标准”）”保存之后的 JPG 文件的数据量较大，逐行显示方式，大多数浏览器都可识别，它是默认的 JPEG 格式。这种格式适合于输出图像。“基线已优化”是在基线算法的基础上将文件进行了优化 Huffman 算法压缩，速度快，保存之后的文件较小。与标准格式相比，两者打印出来的效果没有差异。但是两者所占的内存明显不同。我们一般会选择“基线已优化”来保存图像文件，但是对于类似几百 MB 的 PSD 文件，用此选项来保存文件，PS 易出现无法保存的弹窗，这时可以将品质调低，一般到 8~9 就可以保存了。“连续”格式将在图像下载时，显示图像的一系列逐渐清晰的各版本。它的好处是可以在图像下载完成前看见图像的低分辨率的版本。一般 500 kB 以上的大图片尽量选择“连续”格式。低分辨率预览对于数据量较大的照片在网页上显示也很重要。所以在网页设计中常会选择第 2 或第 3 种。

5.6.2 二值图像压缩标准 JBIG

1980 年，CCITT T.4 建议将文件传真三类机（G3）的一维编码标准 MH，作为二维编码标准的改进相对元地址指定编码（Modified Relative Element Address Designate, MR）。1994 年，CCITT T.6 建议了四类传真机（G4）标准编码方案：二次改进霍夫曼编码（Modified MR, MMR）。二值图像通过 MH，

MR 和 MMR 等典型的编码方法,已在传真机等图像通信中得到广泛应用。将一幅二维图像按位平面进行分解,可以得到若干二值图像。例如,一幅灰度为 256 级的图像可以被分解为 8 幅二值图像。因此,通过二值图像编码可以对灰度图像进行分层传送。支持分层图像传送的编码方法的基本思路是:首先传送过去一幅分辨率较低的概要图像(outline image),然后随传送数据的不断到来,所得到的图像质量逐步提高。

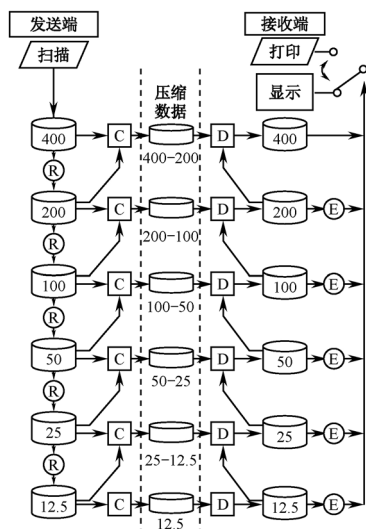


图 5.17 图像的分层传送原理

JBIG (Joint Bi-level Image Coding Experts Group) 所提出的标准化图像正是上面所叙述的分层传送图像,其原理如图 5.17 所示。图中的数字表示像素/英寸,即 dpi。R 表示分辨率递归分层过程,C 表示压缩过程,D 表示解压缩过程,E 表示在接收端按一定的分辨率再现图像。原图像的分辨率设为 400 dpi 或更高一些。

图像在传送时,发送端递归地将图像在水平、垂直两个方向上按分辨率的 1/2 (half-resolution) 求出各层概要图像,直至最低分辨率。假设最低分辨率为 25 dpi,则此时的图像仅相当于原图像的 $(1/16) \times (1/16)$,很适合作为图元使用。

在传送时,首先将分辨率为最低层的图像进行编码、传送。接着,将水平和垂直方向上分辨率为前一个被发送的图像两倍的图像编码、传送。如此下去,直至最上层的 1/2 分辨率的图像。这样,接收端从得到第一幅缩小图像开始,就得到了一个被传送图像的整体概要。随着分辨率不断增大,直至 1/2 图像的到来,所得到的灰度图像也越来越清晰,直至完全恢复为原图像。在这个过程中,如果用户认为所传送的图像不是自己所需要的,或者已经足够清楚,不必进一步细化了,就可以通过一个中断指示,停止传送。

随着通信技术的发展,新一代通信服务内容中,图像数据的传输和检索成为研究的重点之一。上述方案更适合于图像检索的要求。因为远地图像检索操作所得的检索结果并不一定是用户所需要的,因而像上述那样分层,逐步将结果传送给用户是最为合理的,用户一旦发现图像是非所要的,可以立即跳过此幅图像。

随着通信技术的发展,新一代通信服务内容中,图像数据的传输和检索成为研究的重点之一。上述方案更适合于图像检索的要求。因为远地图像检索操作所得的检索结果并不一定是用户所需要的,因而像上述那样分层,逐步将结果传送给用户是最为合理的,用户一旦发现图像是非所要的,可以立即跳过此幅图像。

JBIG 所讨论的标准化内容,不仅包括了编码方式,而且还包括了作为图像缩小操作的前处理规则的描述,这为今后方法的进步,提供了一个开放式的环境,从而为不论通过何种缩小方式都能进行编码和传送提供了可能性。这样,当有更好的缩小方法被提出后,可以吸收到方案中。

JBIG 规定编码方法必须满足的条件和对之进行评价的项目如下。

- (1) 无损编码 (Reversibility or Lossless Coding)。
- (2) 即使接收端不具有帧存储器,系统也能在顺序传送模式下正常工作。
- (3) 编码和解码操作是实时的。
- (4) 压缩和恢复两个功能在时间和复杂性方面是对称的。

(5) 具有比 MMR 更高的压缩能力。与 MMR 相比,JBIG 算法对于打印字符的扫描图像的 C_r 是其 1.1~1.5 倍;计算机打印的字符图像的 C_r 是其 5 倍;由抖动 (dithering) 或半调 (half toning) 表现出来的灰度图像的 C_r 是其 2~30 倍。

(6) 利用同一个数据库可以同时支持顺序和逐层两种压缩传送模式。通过参数定义,采用 PRES (Progressive Reduction Standard) 算法实现二值图像的累进编码。这种编码方式获得的图像的主观质量明显地优于简单地亚采样后的图像。

(7) 不允许使用全帧预扫描 (单路执行算法)。

(8) 在 64 kbps 传输速率下能够做到解码。

(9) 鲁棒性。不论是对于一般图像、计算机生成的图像, 还是对伪灰度、伪彩色图像等, 都能做到有较高的编码压缩能力。

5.6.3 JPEG2000 静态图像压缩标准

目前, 基于整数小波变换的静态图像压缩编码在图像领域取得了较大的成功, 特别是新的静态图像压缩标准 JPEG2000 中采用了基于提升方法的整数小波变换。JPEG2000 图像文件扩展名可为 jpf、jp2 和 j2k 等。

1. JPEG2000 标准制定的目的

JPEG 标准在短短的几年内就获得极大的成功, 目前网站上大部分图像采用的都是 JPEG 压缩标准。然而, 随着多媒体应用领域的进一步扩展, 传统 JPEG 压缩技术在高压压缩率和更多的新功能方面已无法满足人们对多媒体图像资料的要求。因此, 国际标准化组织 (ISO) 制定了新一代静止图像压缩标准 ISO 15444, 即 JPEG2000。

JPEG2000 标准的目标是进一步改进目前压缩算法的性能, 以适应低带宽、高噪声的环境, 以及医疗图像、电子图书馆、传真、互联网上服务和保安等方面的应用。JPEG2000 还将彩色静态画面采用的 JPEG 编码方式与二值图像采用的 JBIG 编码方式统一起来, 成为对应各种图像的通用编码方式。由于 JPEG2000 的特性及功能显著, 且支持旧版本的标准, 因此在需要有较好的图像质量、较低的比特率或是一些特殊特性的要求 (渐进传输和感兴趣区域编码等) 时, JPEG2000 将是最好的选择。但在一些低复杂度的应用中, 并不要求 JPEG2000 完全代替 JPEG, 因为 JPEG2000 的算法复杂度往往不能满足这些领域的要求。目前, JPEG2000 的使用存在版权和专利的风险。

JPEG2000 的制定应用的是导向方式, 即先提出应用, 再制定规格。JPEG 2000 的应用领域可大致分成两个方面: 一方面, 面向传统的 JPEG 市场, 如打印、扫描、数字摄像、遥感等; 另一方面, 面向一些新兴的应用领域, 如网络传输、彩色传真、无线通信、医疗影像、电子商务等。虽然 JPEG2000 在技术上有一定的优势, 但是目前大多数的浏览器仍然没有内置支持 JPEG2000 图像文件的显示功能。

2. JPEG2000 标准提供的主要特征

JPEG2000 标准提供了一套新的特征, 这些特征对于一些新产品 (如数码相机) 和应用 (如互联网) 是非常重要的。它把 JPEG 的顺序模式、渐进模式、无损模式和分层模式四种模式集成在一个标准之中。JPEG2000 与传统 JPEG 最大的不同, 在于它放弃了 JPEG 所采用的以离散余弦变换为主的区块编码方式, 而采用以小波变换为主的多解析编码方式。JPEG2000 标准中无损压缩和有损压缩所采用的小波分别是基于提升方案的 (5, 3) 整数小波和 Daubechies (9, 7) 整数小波。编码端以最大的压缩质量和最大的图像分辨率压缩图像, 在解码端可以从码流中以任意的图像质量和分辨率解压图像, 最大可达到编码时的图像质量和分辨率。

JPEG2000 的最主要的特征如下。

(1) 高压压缩率: 由于在离散小波变换算法中, 图像可以转换成一系列可更加有效存储像素模块的“小波”, 因此, JPEG2000 格式的图片压缩比可在现在的 JPEG 基础上再提高 10%~30%, 而且压缩后的图像显得更加细腻平滑, 这一特征在互联网和遥感等图像传输领域有着广泛的应用。当然, 就图像整体压缩性能来说, 在压缩比小于 10:1 的低压压缩比情形下, 传统的 JPEG 图像质量有可能要比 JPEG2000 要好。但在压缩比较高情形下, JPEG2000 显现优势。

(2) 无损压缩和有损压缩: JPEG2000 提供无损和有损两种压缩方式, 无损压缩在许多领域是必需的, 且 JPEG2000 在无损压缩下仍然有较高的压缩率。例如医学图像和重要的图像档案通常是不能忍受有损压缩, 所以 JPEG2000 在图像品质要求比较高的图像的分析 and 处理中已经得到较好的应用。

同时, JPEG2000 提供的是嵌入式码流, 允许从有损到无损的渐进解压。在有损压缩下, JPEG2000 一个明显的优点是没有 JPEG 压缩中的马赛克效应, 但传统的 JPEG 压缩一样, JPEG2000 也存在由于高频分量的损失而导致的模糊失真现象。

(3) 渐进传输: 现在网络上的 JPEG 图像下载时是按“块”传输的, 因此只能一行一行地显示, 而采用 JPEG2000 格式的图像支持渐进传输 (Progressive Transmission)。这一特性有利于用户对不同分辨率图像的传输、打印和保存。

(4) 感兴趣区域压缩: 可以指定图片上的感兴趣区域 (Region Of Interest, ROI), 然后在压缩时对这些区域指定压缩质量, 或在恢复时指定某些区域的解压缩要求。这是因为小波在空间和频率域上具有局域性, 要完全恢复图像中的某个局部, 并不需要所有编码都被精确保留, 只要对应它的一部分编码没有误差即可。

(5) 码流的随机访问和处理: 这一特征允许用户在图像中随机地定义感兴趣区域, 使得这一区域的图像质量高于其他图像区域; 码流的随机处理允许用户进行旋转、移动、滤波和特征提取等操作。

(6) 容错性: 在码流中提供容错性有时是必要的, 例如在无线等传输误码很高的通信信道中传输图像时, 没有容错性是让人不能接受的。

(7) 开放的框架结构: 为了在不同的图像类型和应用领域优化编码系统, 提供一个开放的框架结构是必需的, 在这种开放的结构中, 编码器只实现核心的工具算法和码流的解析, 如果需要解码器, 可以要求数据源发送未知的工具算法。

(8) 基于内容的描述: 图像文档、图像索引和搜索在图像处理中是一个重要的领域, MPEG-7 就是支持用户对其感兴趣的各类“资料”进行快速、有效的检索的一个国际标准。基于内容的描述在 JPEG2000 中是压缩系统的特性之一。

3. JPEG2000 的基本框架和实现

参考图 5.10 所示的变换编码系统原理框图。首先对原始图像数据进行离散小波变换 (DWT), 然后对变换后的小波系数进行量化, 接着对量化后的数据熵编码, 最后形成输出码流。解码器是编码器的逆过程, 首先对码流进行熵解码, 然后解量化和小波逆变换, 最后生成重建图像数据。

与 JPEG 相似, JPEG2000 把原图像分解成各个成分 (亮度信号和色度信号), 处理对象不是整幅图像, 而是把图像分成若干矩形图像片 (image tiles), 图像片是原始图像和重建图像的基本处理单元, 对每一个图像片进行独立的编解码操作。在对每个图像片进行小波变换之前, 通过减去一个相同的数量值对所有的图像片进行水平移位。对分解后的小波系数进行量化并组成矩形的编码块 (code-block)。对编码块中的系数“位平面”熵编码。为了使码流具有容错性, 在码流中添加相应的标识符 (Marker)。编码器的最后使用了算术编码器, 在 JPEG2000 中使用的是 MQ 编码器, MQ 编码器在本质上与 JPEG 中的 QM 编码器很相似。

一般在压缩比达到 100:1 的情形下, 采用 JPEG 压缩的图像已经严重失真并开始难以识别了, 但 JPEG2000 的图像仍可识别。

5.7 实验: 图像编码与压缩

1. 实验目的

- (1) 了解图像的压缩编码原理。
- (2) 掌握常用的图像压缩算法。

2. 实验主要仪器设备

(1) 台式机或笔记本电脑。

(2) 数码单反相机。

(3) MATLAB 软件（含图像处理工具箱）和高版本 Photoshop（含 Camera Raw 插件，以解释相机原始数据文件，该插件使用有关相机的信息以及图像元数据来构建和处理彩色图像）。

3. 实验原理

(1) 去除数据冗余度可以有效地压缩数据。

(2) 图像编码压缩的主要技术指标：压缩比、客观评价指标（MSE、PSNR 等）、主观评价。

4. 实验内容

(1) 利用变换编码压缩图像信息。

(2) 计算压缩算法的性能。

(3) 正确选择数码相机中 JPEG 的画质。比较 RAW 记录画质与 JPEG 格式画质，比较不同压缩比的 JPEG 画质。

(4) 采用 JPEG2000 存储和显示图像，并与 JPEG 图像的数据量和质量进行比较。

5. 实验步骤

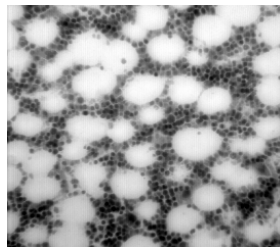
(1) 选择图 5.18 中的原始图像作为测试图像。



(a) Cameraman.tif



(b) westconcordorthoto.tif



(c) bonemarr.tif

图 5.18 供编码压缩的测试图像

(2) 根据设计要求选择图像变换编码与压缩方法。

(3) 设计编码与压缩算法。

(4) 显示原始图像和压缩重构图像。

(5) 计算压缩的性能指标：压缩率和保真度。

(6) 画出压缩率和保真度的关系曲线。

(7) 用数码单反相机拍摄你感兴趣的景象，选择 RAW 和 JPEG 文件同时存储，比较 RAW 和 JPEG 文件的数据量、不同画质 JPEG 存储的压缩比。观测用 Photoshop 打开 RAW 展示的 Camera Raw 版本号 and 相机机型、相关成像参数，如光圈、速度、感光度、焦距、白平衡、色温、色调等。你可以重新调整有关参数以得到所需的效果。在调整过程中，原来的相机原始数据将保存下来。调整内容将作为元数据存储于附带的附属文件、数据库或 DNG 格式文件中。

(8) 打开高质量的 TIFF 图像，利用高版本 Photoshop 分别用 JPEG 和 JPEG2000 格式进行存储和显示图像，对两者的数据量和质量做比较。

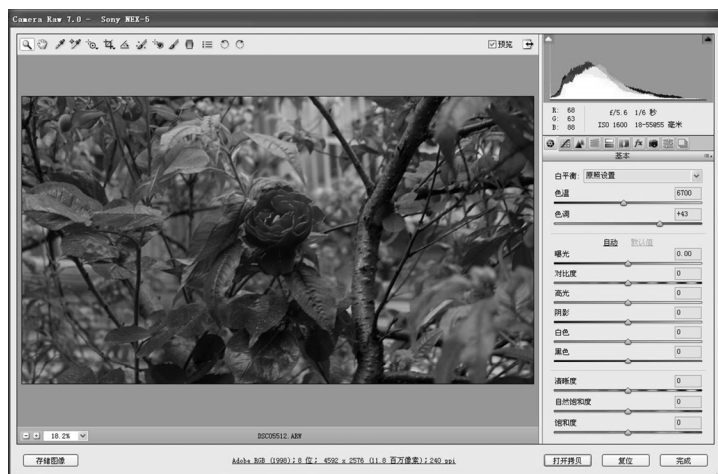


图 5.18 Photoshop Camera Raw 界面

6. 实验报告要求

- (1) 说明压缩原理与算法。
- (2) 原始图像和压缩重构图像。
- (3) 计算压缩性能, 说明为什么同一压缩算法对不同的图像压缩效果不同。
- (4) 回答思考题。
- (5) 心得和体会。

7. 预习要求

- (1) 了解图像编码压缩的基本原理和算法。
- (2) 掌握 MATLAB 图像处理工具箱中有关图像编码压缩的函数。
- (3) 画出所采用算法的流程框图, 编写相应的程序。

8. 思考题

- (1) DFT 与 DCT 的性能比较。
- (2) 讨论压缩率和保真度的关系。
- (3) 讨论同一算法对不同原始图像压缩性能的影响。
- (4) 数码单反相机中的 JPEG 记录画质和 RAW 记录画质有什么区别?

本章小结

数字视频编码压缩技术是现代数字视频技术和多媒体技术的重要组成部分。本章介绍的主要内容有:

(1) 数据压缩的基本概念: 编码冗余。像素间的相关性形成的冗余, 视觉特性和显示设备引起的冗余, 使图像编码压缩成为可能。

(2) 数据压缩方法有: 无损压缩有霍夫曼编码、香农编码、游程编码、算术编码和轮廓编码等。有损压缩有预测编码、变换编码。混合编码将预测编码与变换编码相结合。

(3) 预测编码的基本原理是利用当前像素与邻近像素的位置关系计算预测信号, 系统所传递的是误差信号, 这种差值的概率分布集中在小数值上, 大差值的概率极小, 有利于用可变码长编码以减少

传送的信息量。按利用的邻近像素的不同,可以是前值预测、一维预测、二维预测、三维预测等。

(4) 变换编码的基本原理是利用坐标变换,如果选择的变换坐标与图像特征相匹配,就可以大大压缩二维数据。重要的变换编码方法是离散余弦变换(DCT),它有快速算法,二维变换可以转化为两次一维变换。DCT变换后的系数相对集中,还可以进一步量化,从而更大幅度地压缩图像的数据量。

(5) 霍夫曼编码、算术编码、游程编码等无损压缩可以降低编码的冗余,在不减少传送信息量的条件下减少了码率。

(6) JPEG标准适用于静止图像。它将图像分解为 8×8 的样值子块,用DCT进行变换、量化、Z字形重排,用霍夫曼码对量化系数进行编码,进一步压缩数据量。

(7) JPEG2000标准是基于DWT的新一代图像压缩标准,同时支持无损压缩和有损压缩。与JPEG相比,JPEG2000在无损压缩下仍然有比较好的压缩率。在有损压缩下无马赛克现象。

理解和掌握数字视频数据压缩的理论基础是本章的重点。熟知各种国际标准的特点、应用领域。了解各种编/解码器的原理和软件和硬件实现。注意,数字视频信号源编码技术的最新发展和应用前景。

思考题与习题

5.1 从哪些方面说明数据压缩的必要性?

5.2 数据没有冗余度能否压缩?为什么?

5.3 如何衡量图像编码压缩方法的性能?

5.4 通常用以未进行压缩的量化比特数 8 bit/像素 为基础讨论压缩比,经过压缩后,平均每个像素的平均比特数为 \bar{r}_c ,求压缩比 C_r 。

5.5 一图像的大小为 640×480 ,256色。用软件工具将其分别转换成24位色BMP,24位色JPEG,GIF(只能转换成256色)压缩格式,24位色TIFF压缩格式,24位色TGA压缩格式,得到的文件大小分别为921 654字节、17 707字节、177 152字节、923 044字节和768 136字节。分别计算每种压缩图像的压缩比。

5.6 讨论图像压缩方法的分类及其各自的特点。

5.7 大部分视频压缩方法是有损压缩还是无损压缩?为什么?

5.8 若图像上任意两像素点的亮度电平值相等,或者任意两时刻同一位置上的像素的亮度电平值相等,能够说明上述两种情况下像素相关吗?为什么?

5.9 根据JPEG算法说明JPEG图像显示时会出现“马赛克”现象的原因。

5.10 讨论混合编码的优点。

5.11 有了离散傅里叶及其快速算法FFT,为什么还要提出离散余弦算法DCT及其快速算法?为什么许多视频国际标准将DCT作为帧内编码的基本压缩算法?

5.12 为什么二维DCT可以转换成两次一维DCT计算?

5.13 图像编码中,采用霍夫曼编码有何优缺点?

5.14 霍夫曼编码是最佳编码,在图像编码中,为什么还要研究算术编码等其他熵编码算法?

5.15 算术编码有何优点?举例说明其适用范围。

5.16 用JPEG标准,对于 $576 \text{ 行} \times 720 \text{ 列}$ 的CCIR601建议分辨率的彩色图像,其亮度分量可分割成多少个子块?而两个色差分量可分别分割成多少子块?

5.17 JPEG为什么要进行彩色空间转换?

5.18 JPEG的量化表有何作用?

5.19 JPEG算法中DCT系数采用Z字形重排有何作用?

第6章 图像增强

➡ 内容提要

图像在成像、采集、传输、处理等过程中不可避免地会造成某些降质。

图像增强是一类对图像降质进行改善的方法，它们有选择性地突出图像的边缘、轮廓、对比度等特征，以便于显示、观察或进一步分析与处理。

图像增强的目的在于采用一系列技术改善图像的视觉效果，提高图像的可懂度；或者将图像转化成一种更适合于人或机器进行分析处理的形式。

图像复原是与图像增强类似的技术，不过其主要目的是为了_{提高图像的保真度}，这将在第7章中介绍。

➡ 内容提要

- 单点增强方法：灰度级校正、灰度变换、灰度直方图变换等。
- 区域增强的平滑方法：邻域平均法、中值滤波和各种边界保持类滤波方法。
- 区域增强的锐化方法：梯度锐化法、拉普拉斯算子、高通滤波及其他常用的锐化算子。
- MATLAB 和 Photoshop 中图像增强的实现方法。

➡ 教学建议

- 本章教学安排6学时。
- 学习本章的先修知识包括高等数学、矩阵的表示与运算、概率论、图像的描述方法、图像变换等。
- 重点让读者了解图像增强方面的内容，包括常见的图像灰度变换、去噪和图像锐化等基本增强方法。

6.1 概述

6.1.1 图像增强的目的

图像在成像、采集、传输、处理等过程中不可避免地会造成某些降质（退化）。如在成像过程中由于光学系统会导致图像几何失真，不同的光照条件会使图像的曝光度差异很大，运动状态下成像会使图像模糊；而在传输过程中，各种噪声和干扰将污染图像。因此，通常需要对降质的图像进行预处理，以满足后期处理或分析的需要。图像复原是改善图像的一类方法，这类方法会尽可能还原图像的本来面目，追求提高图像的保真度（fidelity）。而图像增强则是一类追求图像可懂度（intelligibility）的方法，通过处理有选择地突出某些感兴趣的信息，便于人或机器分析这些信息，抑制一些无用的信息，以提高图像的使用价值。

图像增强（image enhancement）是指对图像的某些特征如边缘、轮廓、对比度等进行强调，以便于显示、观察或进一步分析与处理。

图像增强处理后的图像比原始图像更适合于具体应用。由于成像和传输等过程的差异性很大，决定了图像增强的方法只能有选择性地使用。实际应用时，需要根据所处理的对象、待解决的问题以及最终要达到的效果等情况，选择合适的图像增强算法，并做适当的优化。

增强的结果主要靠人的主观感觉加以评价，很难对增强结果加以定量描述。这既给增强算法的设计带来较大的自由度，也给不同增强算法的优劣比较带来了困难。

6.1.2 图像增强技术的分类

图像增强技术大致分为空间域增强和频率域增强两类。

1. 空间域增强法

在空间域直接对像素灰度值进行运算，如图 6.1 所示。图中， $f(x, y)$ 是待增强的原始图像， $g(x, y)$ 是已增强的图像， $h(x, y)$ 是空间运算函数。显然对灰度变换、直方图变换等点操作有

$$g(x, y) = f(x, y) \cdot h(x, y) \quad (6.1)$$

而对于区域操作（如平滑、锐化等）有

$$g(x, y) = f(x, y) * h(x, y) \quad (6.2)$$

式中，* 表示卷积运算。

2. 频率域增强法

先对待增强的图像进行变换 T （如 DFT 或 DWT），即 $f(x, y) \rightarrow F(u, v)$ ，在频率域利用二维滤波器 $H(u, v)$ 对 $f(x, y)$ 进行滤波，得到新的频谱 $G(u, v)$ ，即

$$G(u, v) = F(u, v) \cdot H(u, v) \quad (6.3)$$

$H(u, v)$ 的性质可能是低通，起平滑作用；也可能是高通，起锐化作用。 $G(u, v)$ 经过逆变换 T^{-1} 即得到增强后的图像 $g(x, y)$ ，也就是 $G(u, v) \rightarrow g(x, y)$ ，如图 6.2 所示。

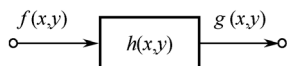


图 6.1 空间域增强模型

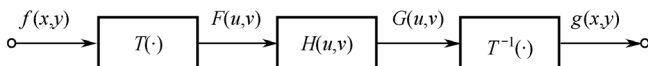


图 6.2 频率域增强模型

实际的图像增强方案可能综合上述两种技术。如同态滤波增强包含了空间域灰度的非线性运算，也有高频增强环节。

本章介绍的是灰度图像增强的基本理论和方法。在第 10 章中，我们还将进一步学习彩色图像的增强技术。

6.1.3 直方图的概念

对于连续图像，其灰度分布的统计特性用概率密度函数（Probability Density Function, PDF）刻画。相应地，离散图像用直方图表示该图像中各种不同灰度级像素出现的相对频率。在数字图像处理中，灰度直方图（gray histogram）是简单且实用的工具，对图像的采集、处理和分析都可以有效地利用它。

灰度直方图的横坐标是图像的各灰度级，纵坐标是各个灰度出现的像素个数。计算并用 n 个灰度级显示图像 I 灰度直方图的 MATLAB 函数是 `imhist(I,n)`。进一步定义归一化的直方图为灰度级出现的相对频率 $P_r(k)$ 。即

$$P_r(k) = n_k / N \quad (6.4)$$

式中， N 表示像素的总数； n_k 表示灰度级为 k 的像素数目。

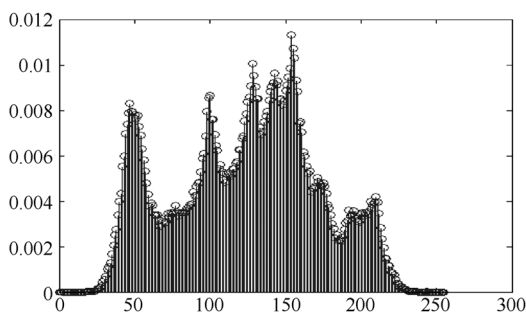
【例 6.1】 求 lena 图像的归一化直方图（如图 6.3 所示）。

【解】 lena 图像是彩色图像，故需进行格式转换。实现的程序如下：

```
I = imread('lena.jpg');
J = rgb2gray(I);          %将彩色图像转换为灰度图像
imshow(J);
N = numel(J);             %求图像像素的总数
Pr = imhist(J)/N;         %显示原始图像的归一化直方图
k=0:255;
figure, stem(k,Pr)
```



(a) lena 图像



(b) lena 图像的直方图

图 6.3 直方图

由图 6.3 可见，灰度直方图（gray-level histogram）反映了数字图像中每一灰度级与其出现频率间的统计关系。它能描述该图像的概貌，如图像的灰度范围、每个灰度级出现的频率、灰度级的分布、整幅图像的平均明暗和对比度等，对图像的采集和进一步处理提供了重要依据。部分数码相机已采用直方图辅助图像采集。Photoshop 中也可方便得到图像的灰度和彩色直方图。对于灰度图像，计算和显示灰度直方图即可。对于彩色图像，分为 R、G、B 三通道和明度（即亮度）分别计算与显示直方图。如图 6.4 所示，选择“窗口”菜单的“直方图”命令（Photoshop CS 版本界面，而 CS 以前的版本是通过“图像”菜单的“直方图”命令来察看，如图 6.5 所示）。“通道”可选项为：RGB

(统计图像的所有 RGB 数据)、红 (统计 R 基色数据)、绿 (统计 G 基色数据)、蓝 (统计 B 基色数据)、明度 (统计灰度数据)、颜色 (将红、绿、蓝和 RGB 通道的统计数据绘制在同一幅图中)。横轴标示亮度值 (0~255), 纵轴标示每种像素的数量。“像素” (Pixels) 为图像的像素总数。“百分位”是从最左边到鼠标指针位置的所有像素数量与图像像素总数之比值。“色阶” (即亮度值, Level)、“数量” (Count)、“百分位” (Percentile) 这三项根据鼠标指针的位置来显示横坐标当前位置的统计数据。

下面讨论“高速缓存级别” (Cache level) 的作用。在处理一些超大的图片 (原图的缓存级别为 1) 时, 会因为像素数量过多而导致每一步操作都很慢。为了提高图片的显示与刷新速度, Photoshop 利用图片的缩小版本来进行计算并显示。当载入一幅图片时, Photoshop 会自动生成许多这幅图片的缩小版本放在临时文件中, 这些缩小图就是高速缓存。缩小比例为 50% (缓存级别为 2)、25% (缓存比例为 3)、12.5% (缓存比例为 4), 以此类推。

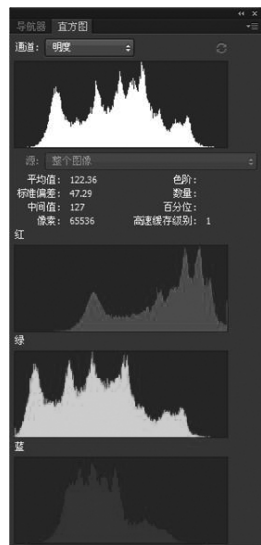


图 6.4 Photoshop 直方图

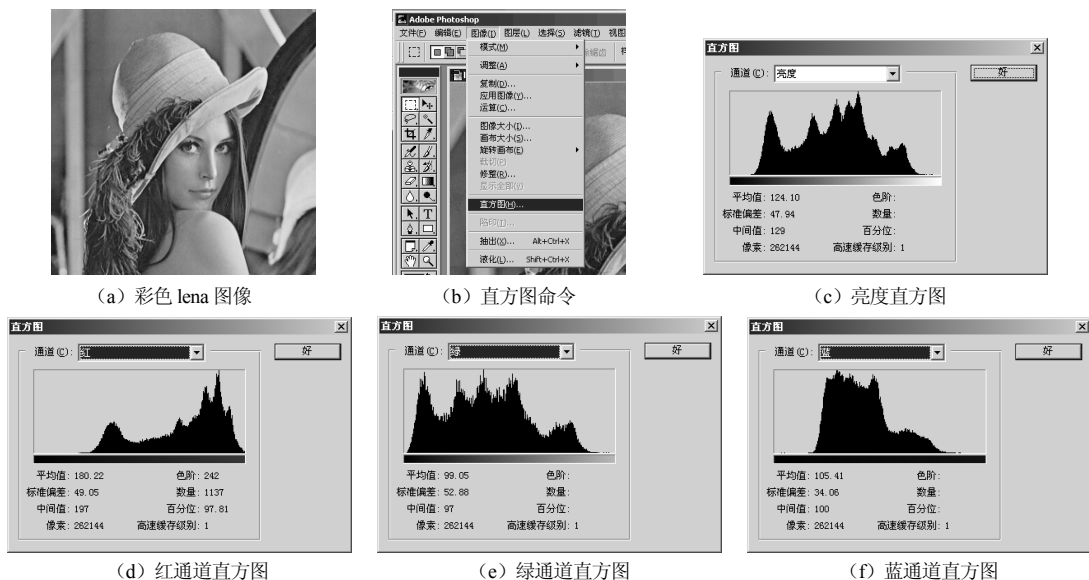


图 6.5 彩色 lena 图像的直方图计算与显示

6.2 灰度修正

图像的退化会引起灰度级的变化。通过简单和有效的点运算, 可以改善图像的显示效果, 达到灰度修正的目的。点运算 (point operation) 是指原始图像的像素灰度值通过运算后产生新图像的对应的灰度值。可见一幅输出图像上每个像素的灰度值仅由相应输入像素的灰度值决定, 而与像素点所在的位置无关, 与相邻的像素之间也没有运算关系。

灰度修正 (gray-level correction) 是图像数字化、图像处理和图像显示软件的重要组成部分。一般有三种方法: (1) 灰度级校正解决成像不均匀问题。(2) 对比度增强解决图像曝光不正确问题。(3) 直方图修正以突出所需要的图像特征。

6.2.1 灰度级校正

在成像过程中，很多因素会导致图像的灰度级发生失真，如光照的强弱、光电转换系统中感光部件的灵敏度、电子元器件特性的不稳定等，均可引起图像亮度分布的不均匀。灰度级校正希望在图像采集系统中对图像像素进行逐点修正，使得整幅图像能够补偿成均匀成像。

设原始图像为 $f(x, y)$ ，实际获得的含噪声的图像为 $g(x, y)$ ，则有

$$g(x, y) = e(x, y) \cdot f(x, y) \quad (6.5)$$

式中， $e(x, y)$ 是具有降质性质的函数。显然，只要知道了 $e(x, y)$ ，就可以重建原始图像 $f(x, y)$ 。然而 $e(x, y)$ 往往未知，需要根据图像降质系统的特性计算或测量。

系统降质函数 $e(x, y)$ 可以简单地采用一幅灰度级全部为常数 C 的图像成像来标定。若经成像系统的实际输出为 $g_C(x, y)$ ，则

$$g_C(x, y) = C \cdot e(x, y) \quad (6.6)$$

根据式 (6.5)，可求得 $e(x, y)$ 为

$$e(x, y) = \frac{g_C(x, y)}{C} \quad (6.7)$$

将上式代入式 (6.5)，就可得实际图像 $g(x, y)$ 经校正后所恢复的原始图像 $f(x, y)$ ，即

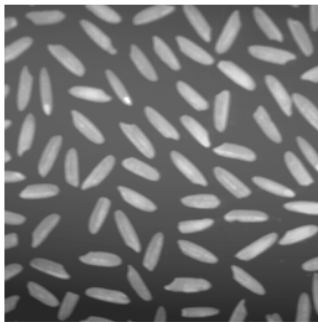
$$f(x, y) = C \cdot \frac{g(x, y)}{g_C(x, y)} \quad (6.8)$$

在实际应用中应当注意两个问题：

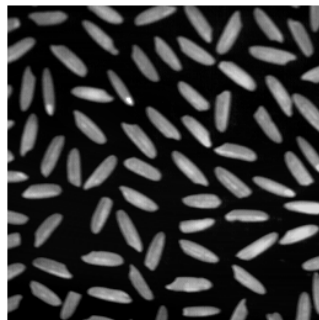
(1) 按式 (6.8) 校正的图像由于乘了系数的影响，有可能出现“溢出”(overflow)现象，即灰度级值可能超过某些记录器件或显示设备输入信号的动态范围，因此需再做适当的灰度级变换，下一节介绍的方法将有效地改变灰度级的范围。

(2) 经灰度级校正后的图像灰度值不一定在原降质图像的量化值上，因此必须对变换后的图像重新进行量化。

图6.6所示为非均匀光照的校正实例。图6.6(a)是一幅由于光照不均匀造成上半部分较下半部分明亮的图像。经灰度级校正后，图6.6(b)所示图像的光照均匀性得到了明显改善。对光照不均匀处理的直接方法是从原图中求出背景的变化，从原图中减去背景图像即可得到较为理想的图像。一种对背景图像估计的方法认为背景是相对缓慢变化的，通过分块操作函数blkproc求出不同块的背景。



(a) 校正前的图像



(b) 校正后的图像

图 6.6 非均匀光照的校正

6.2.2 灰度变换

在一些应用场合,为了将图像灰度级的整个范围或一段范围扩展或压缩到记录或显示设备的动态范围内,从而使图像变得更加清晰、图像上的特征更加明显,可以采用灰度变换方法。

灰度变换(gray-level transformation)可使图像动态范围增大,图像对比度扩展,是图像增强的重要手段之一。例如,在处理数码相片时,有时可能因为环境光源太暗,使灰度值偏小,就会使图像太暗看不清。如果环境光源太亮,又使图像泛白。通过灰度变换,就可以将灰度值调整到合适的程度,使处理后的相片变得悦目。

灰度变换可分为线性变换、分段线性变换和非线性变换几种方法。

1. 线性变换(linear transformation)

设原始图像像素灰度 f 的范围为 $[a, b]$,线性变换后图像像素灰度 g 的范围为 $[a', b']$,如图6.7所示。

灰度 g 与灰度 f 之间的关系为

$$g = a' + \frac{b' - a'}{b - a}(f - a) \quad (6.9)$$

可见,如果 $b' - a' > b - a$,则线性变换使得图像灰度范围增大,即对比度增大,图像会变得更加清晰;如果 $b' - a' < b - a$,则使得图像灰度范围缩小,即对比度减小。在曝光不足或过度的情况下,图像灰度可能会局限在一个很小的范围内,这时在显示器上看到的将是一个模糊不清、没有灰度层次的图像。采用线性变换对图像的每一个像素灰度做线性拉伸,将有效地改善图像的视觉效果。以曝光不足为例,选取 $b' - a' > b - a$,则使曝光不充分的图像中黑的更黑,白的更白,从而有效地提高图像灰度的对比度。MATLAB用于对比度调整的函数为`imadjust`,其主要形式为

$$J = \text{imadjust}(I, [\text{low high}], [\text{bottom top}], \text{gamma})$$

此函数将灰度图像 I 转换为图像 J ,使值从 low 到 high 与从 bottom 到 top 相匹配,值大于 high 或小于 low 的被剪去,即小于 low 的值与 bottom 相匹配,大于 high 的值与 top 相匹配。 gamma (γ ,伽马,用来衡量输入和输出幂函数转换特性的指数)用来指定描述 I 和 J 值关系曲线的形状,默认时 $\text{gamma} = 1$,表示线性变换。

【例 6.2】采用线性变换进行图像增强。应用 MATLAB 函数 `imadjust` 将图像在 $0.3 \times 255 \sim 0.7 \times 255$ 灰度之间的值通过线性变换映射到 $0 \sim 255$ 之间。

【解】在式(6.9)中各值分别取 $a = 0.3 \times 255$, $b = 0.7 \times 255$, $a' = 0$, $b' = 255$ 。实现的程序如下:

```
I = imread('pout.tif');
imshow(I);
figure, imhist(I); %显示原始图像的直方图
J = imadjust(I, [0.3 0.7], []); %使用 imadjust 函数进行灰度的线性变换
figure, imshow(J);
figure, imhist(J) %显示变换后图像的直方图
```

程序运行的结果如图6.8所示。

如图6.9所示,对图1.17所示的显微图像进行灰度变换增强。图6.9(a)中是灰度图像,图6.9(b)是其直方图,由以上二图可见,原图的灰度图像较模糊、不很清晰,直方图局限在 $[25, 200]$ 左右。对灰度范围进行线性拉伸,使图像灰度值映射到 $[0, 255]$ 之间,如图6.9(d),图像视觉效果得到改善,如图6.9(c)。

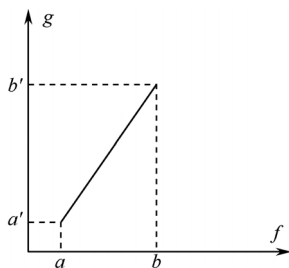


图 6.7 线性变换

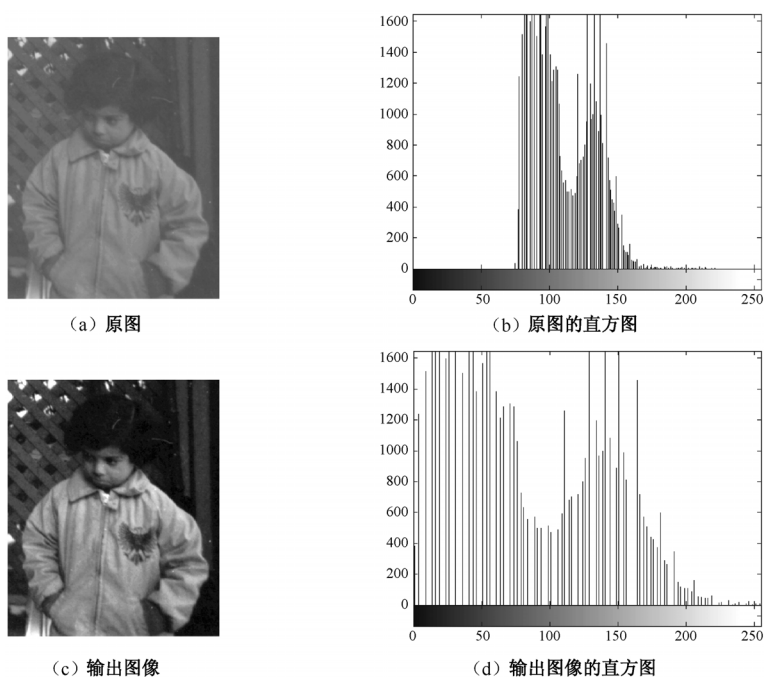


图 6.8 灰度线性变换

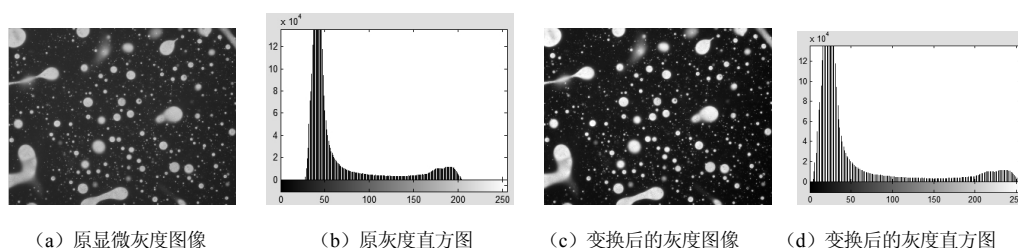


图 6.9 显微图像灰度变换

2. 分段线性变换 (piecewise linear transformation)

对整个灰度区间进行分段, 采用分段线性函数进行变换。这种变换突出了感兴趣的目标或灰度区间, 相对抑制那些不感兴趣的灰度区间。常用的是三段线性变换, 如图 6.10 所示。

对应的数学表达式为

$$g = \begin{cases} (c/a)f, & 0 \leq f < a \\ [(d-c)/(b-a)][f-a] + c, & a \leq f < b \\ [(g_{\max}-d)/(f_{\max}-b)][f-b] + d, & b \leq f \leq f_{\max} \end{cases} \quad (6.10)$$

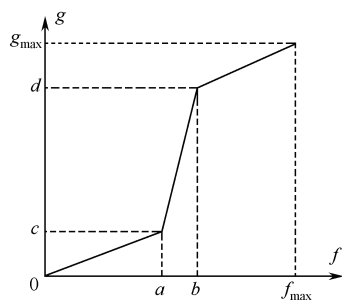


图 6.10 三段线性变换

图 6.9 中对灰度区间 $[a, b]$ 进行了线性拉伸, 而灰度区间 $[0, a]$ 和 $[b, f_{\max}]$ 则被压缩。仔细调整折线拐点的位置并控制分段直线的斜率, 可以对图像的任一灰度区间进行拉伸或压缩。

在某些应用场合, 可将过黑或过白的灰度级压缩成一个灰度级。这样就指定了一个变换段 $[a, b]$, 只有灰度在 $[a, b]$ 段内的像素才会增强。灰度小于 a 的编程为某个暗色 g_{\min} (可

以设为 0)，而灰度大于 b 的变成某个亮色 g_{\max} （可以设为 255）。这时，式 (6.10) 改写为

$$g = \begin{cases} g_{\min}, & 0 \leq f < a \\ [(d-c)/(b-a)][f-a] + c, & a \leq f < b \\ g_{\max}, & b \leq f \leq f_{\max} \end{cases} \quad (6.11)$$

例如，在遥感图像分类中，感兴趣的地貌特征可能有明显的灰度变化，而那些过黑或过白的像素往往对应于玄武岩、水、冰等。图 6.11 是一个三段线性变换的实例。图 6.11 (a) 为 Couple 原始图像，图 6.11 (b) 为三段式增强后的效果。可见增强后的图像更加明快、清晰。



(a) 原始图像

(b) 增强效果

图 6.11 三段线性变换实例

3. 非线性灰度变换

可以根据需要制定非线性函数对灰度进行变换，典型的有幂函数、对数函数、指数函数等。幂函数变换中典型的有来自视频系统的伽马校正 (gamma (γ) correction) 技术。 γ 失真源于显示器的响应曲线，即其亮度与输入电压的指数为 γ 的非线性关系。在视频系统，摄像过程通过 Gamma 校正对输入视频信号的幅度进行预失真，最终保证整个视频系统的传输特性呈现线性。如图 6.12 所示， $\gamma = 2.2$ ，则摄像端进行 γ 校正，对灰度进行提升，0.218 提升为 $0.218^{1/\gamma} = 0.5$ （图中①），而在显示端 0.5 又还原成 $0.5^\gamma = 0.218$ （图中③）。在上述的灰度“线性变换”中用到的 MATLAB 函数 imadjust 的 gamma 参数正是起到这样的作用。选项 $\text{gamma} = 1/2.2 = 0.45$ ，修改例 6.2 的程序：

```
J = imadjust(I, [0.3 0.7], [], 1/2.2);
```

运行结果如图 6.13 所示，同样可以取得较好的增强效果。Photoshop 在“图像”菜单中“曝光度”命令的“灰度系数校正”选项或“曲线”命令都可以交互式地进行伽马校正。

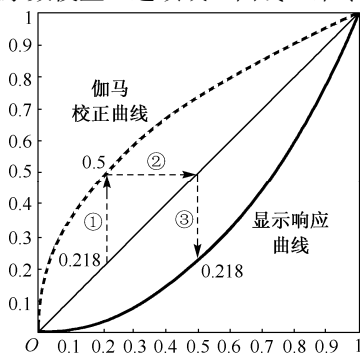
图 6.12 $\gamma = 2.2$ 时的伽马校正原理图

图 6.13 对实际图像进行伽马校正

对数变换可以增强低灰度级的像素，扩展低灰度区，压制高灰度级的像素，当希望对图像的低灰度区进行较大的拉伸而对高灰度区压缩时，可采用这种变换，它能使图像灰度分布与人的视觉特性相匹配。对数变换的一般表达式为



图 6.14 对数变换后的 Couple 图像

$$g = a + c \cdot \lg(f + 1) \quad (6.12)$$

式中, a 是偏置参数, c 是比例系数。图 6.14 所示为对数变换后的 Couple 图像。

指数变换的作用与对数变换相反, 用来压缩低灰度区。指数变换的一般表达式为

$$g = b \cdot \exp[c \cdot (f - a)] \quad (6.13)$$

式中 a, b, c 是变换参数。

由于增强的效果要考虑人的主观感觉, 所以实际增强操作往往要采用交互方式, 而且与所考虑的应用有着密切的联系。

6.2.3 灰度直方图变换

直方图的修正是图像灰度级变换的最常用方法。自然图像由于其灰度分布可能集中在较窄的区间, 引起图像不够清晰。例如, 过曝光图像的灰度级集中在高亮度范围内; 而曝光不足将使灰度级集中在低亮度范围内。采用直方图变换后, 可使图像的灰度间距拉开或使灰度分布均匀, 从而增大对比度, 使图像清晰, 达到增强的目的。

直方图变换通常有直方图均衡化 (histogram equalization) 及直方图规定化 (histogram specification) 两类。本书主要介绍直方图均衡处理。

直方图均衡化是通过在原图像进行某种变换, 使得图像的直方图变为均匀分布的直方图, 从而达到增强的效果。下面先讨论连续图像的均衡化问题, 然后推广到离散图像上。

为讨论方便起见, 以 r 和 s 分别表示归一化了的原图像灰度和经直方图变换后的图像灰度。当 $r = s = 0$ 时, 表示黑色; 当 $r = s = 1$ 时, 表示白色; 即 $r, s \in [0, 1]$, 这个范围表示像素灰度在黑白之间变化。

在 $[0, 1]$ 区间内的任一个 r , 经变换函数 $T(r)$ 都可产生一个 s , 且

$$s = T(r) \quad (6.14)$$

式中, $T(r)$ 应当满足下列条件:

- (1) 在 $0 \leq r \leq 1$ 内为单调递增函数;
- (2) 在 $0 \leq r \leq 1$ 内, 有 $0 \leq T(r) \leq 1$ 。

条件 (1) 保证灰度级从黑到白的次序不变, 条件 (2) 确保映射后的像素灰度在允许的范围内。逆变换关系为

$$r = T^{-1}(s) \quad (6.15)$$

$T^{-1}(s)$ 对 s 同样满足上述两个条件。

由概率论可知, 如果已知随机变量 r 的概率密度为 $p_r(r)$, 而随机变量 s 是 r 的函数, 则 s 的概率密度 $p_s(s)$ 可以由 $p_r(r)$ 求出。假定随机变量 s 的分布函数用 $F_s(s)$ 表示, 根据分布函数的定义有

$$F_s(s) = \int_{-\infty}^s p_s(s) ds = \int_{-\infty}^r p_r(r) dr \quad (6.16)$$

根据概率密度函数是分布函数的导数的关系, 式 (6.16) 两边对 s 求导可得

$$p_s(s) = \frac{d}{ds} \left[\int_{-\infty}^r p_r(r) dr \right] = p_r(r) \frac{dr}{ds} = p_r(r) \frac{d}{ds} [T^{-1}(s)] \quad (6.17)$$

从式 (6.17) 可以看出, 通过变换函数 $T(r)$ 可以控制图像灰度级的概率密度函数, 从而改善图像的灰度层次, 这就是直方图均衡的基础。

从人眼视觉特性来考虑,一幅图像的灰度直方图如果是均匀分布的,即 $p_s(s) = k$ (归一化后 $k=1$) 时,感觉上该图像比较协调。因此要求将原图像进行直方图均衡化,以满足人眼视觉的要求。

因为归一化假定

$$p_s(s) = 1 \quad (6.18)$$

由式 (6.17) 得 $ds = p_r(r) dr$ 。两边积分得

$$s = T(r) = \int_0^r p_r(r) dr \quad (6.19)$$

式 (6.19) 就是所求得变换函数。它表明当变换函数 $T(r)$ 是原图像直方图的累积分布函数时,能达到直方图均衡化的目的。

对于灰度级为离散的数字图像,用频率来代替概率,则变换函数 $T(r_k)$ 的离散形式可表示为

$$s_k = T(r_k) = \sum_{i=0}^k p_r(r_i) = \sum_{i=0}^k \frac{n_i}{n} \quad (6.20)$$

式中 $0 \leq r_k \leq 1, k=0,1,2,\dots,L-1$ 。可见,均衡后各像素的灰度值 s_k 可直接由原图像的直方图算出。

下面通过一个例子说明图像直方图均衡的过程。

【例 6.3】 假定有一幅总像素为 $n=64 \times 64$ 的图像,灰度级数为 8,各灰度级分布列于表 6.1 中。试对其进行直方图均衡化。

表 6.1 一幅图像的灰度级分布

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s'_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1
s''_k	1/7	3/7	5/7	6/7	6/7	1	1	1
s_k	1/7	3/7	5/7		6/7			1
n_{s_k}	790	1023	850		985			448
$p_r(s_k)$	0.19	0.25	0.21		0.24			0.11

【解】

$$(1) \text{ 按式 (6.14) 求变换函数 } s'_k, \quad s'_0 = T(r_0) = \sum_{j=0}^0 p_r(r_j) = 0.19$$

$$s'_1 = T(r_1) = \sum_{j=0}^1 p_r(r_j) = 0.19 + 0.25 = 0.44$$

类似地计算出 $s'_2 = 0.65, s'_3 = 0.81, s'_4 = 0.89, s'_5 = 0.95, s'_6 = 0.98, s'_7 = 1$ 。

(2) 计算 s''_k : 考虑输出图像灰度是等间隔的,且与原图像灰度范围一样取 8 个等级,即要求最终的值 $s_k = k/7, k=0,1,\dots,7$ 。因而需要对 s'_k 进行重新量化后加以修正(采用四舍五入法),得到

$$s''_0 = 1/7, s''_1 = 3/7, s''_2 = 5/7, s''_3 = 6/7, s''_4 = 6/7, s''_5 = 1, s''_6 = 1, s''_7 = 1$$

(3) s_k 的确定: 由 s''_k 可知,输出图像的灰度级仅为 5 个级别,它们是

$$s_0 = 1/7, s_1 = 3/7, s_2 = 5/7, s_4 = 6/7, s_7 = 1$$

(4) 计算对应每个 s_k 的 n_{s_k} : 因为 $r_0 = 0$ 映射到 $s_0 = 1/7$, 所以有 790 个像素在输出图像上变成 s_0 。同样, $r_1 = 1/7$ 映射到 $s_1 = 3/7$, 所以有 1023 个像素取值 $s_1 = 1/7$ 。 $r_2 = 2/7$ 映射到 $s_2 = 5/7$, 因此有 850 个

像素取值 s_2 。又因为 r_3 和 r_4 都映射到 $s_4=6/7$ ，因此有 $656+329=985$ 个像素取值 s_4 。同理有 $245+122+81=488$ 个像素变换到 $s_7=1$ 。

(5) 计算 $p_s(s_k)=n_{s_k}/n$

以上各步计算的结果填在表 6.1 中。

IPT 中的 `histeq` 函数用于直方图均衡，主要形式为 $J = \text{histeq}(I, n)$ ，可将灰度图像 I 转换成具有 n 个离散灰度级的灰度图像 J 。 n 的默认值为 64。

【例 6.4】 在 MATLAB 环境中，采用直方图均衡的方法对 `tire.tif` 图像进行图像增强。

【解】 程序如下：

```
I = imread('tire.tif');
J = histeq(I);           %完成直方图均衡化
imshow(I);               %显示直方图均衡化前的图像
figure, imhist(I);        %均衡化前的直方图
figure, imshow(J);        %显示直方图均衡化后的图像
figure, imhist(J);        %均衡化后的直方图
```

直方图均衡化的结果如图 6.15 所示。图 6.15 (a) 和图 6.15 (b) 分别为一幅灰度级数为 8 bit 的图像及其直方图。原图像较暗，从直方图中可以看出大部分像素的灰度集中在某些灰度值上。图 6.15 (c) 和图 6.15 (d) 分别是经均衡化得到的图像及其直方图。比较图 6.15 (a) 和图 6.15 (c) 可以看出，图 6.15 (c) 的反差增大了，许多细节更加清晰，对应的直方图变得平坦多了。

从图 6.15 中还可以了解到：由于采用离散公式，其概率密度函数是近似的，原直方图上频率较小的某些灰度级被合并到一个或几个灰度级中，频率小的部分被压缩，频率大的部分被增强。从图 6.15 (d) 中可以看到结果是一种近似的、非理想的均衡结果。虽然均衡所得图像的灰度直方图不是很平坦，但从分布来看，比原图像直方图平坦多了，而且动态范围扩大了。

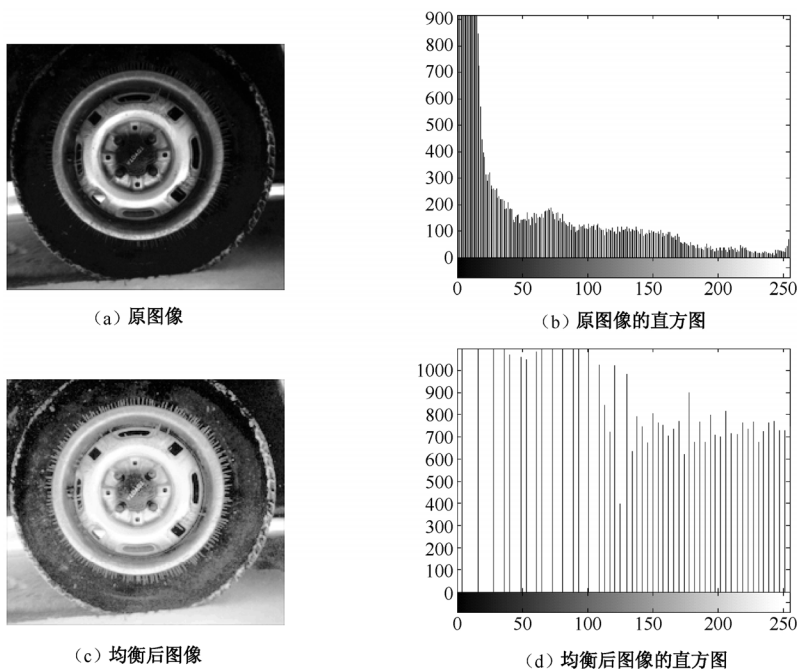


图 6.15 直方图均衡

6.3 同态增晰

6.3.1 问题的由来

在实际应用中，一类图像由于物体受到不均匀的照度，其灰度级动态范围很大，黑与白形成强烈反差，而我们感兴趣目标的灰度级范围却很小，分不清目标的灰度层次和细节，图像上对应照度暗的部分，其细节较难辨别。这时，采用一般的灰度线性变换无法解决以下矛盾状态的增强问题：扩展灰度级虽可以提高目标的反差，但会使动态范围更大；而压缩灰度级，虽可以减少动态范围，但目标的灰度层次和细节就会更看不清。同态滤波的目的是，消除不均匀照度的影响而又不损失图像细节。

同态滤波可以消除不均匀照度的影响，增强图像细节，同时也是图像复原的一种方法。本书第7章将对同态复原做进一步介绍。

6.3.2 增晰原理

图像同态增晰 (homomorphic enhancement) 系统采用合适的滤波特性函数，可以既使图像灰度动态范围压缩，又能让感兴趣的物体图像灰度级扩展，从而使图像清晰。

同态系统是服从广义叠加原理的各类非线性系统。对同态系统信号处理，特别适合处理乘法组合信号和卷积组合信号。图像同态增晰等处理属于前者，其原理框图如图 6.11 所示，而有些语音信号处理、地震信号处理等属于后者。

一般情况下，图像是物体对照明光的反射。自然景物的图像 $f(x, y)$ 是由两个分量乘积所组成的，即照明函数 $i(x, y)$ 和反射函数 $r(x, y)$ 的乘积，

$$f(x, y) = i(x, y) \cdot r(x, y) \quad (6.21)$$

可见，实际图像的灰度由照明分量和反射分量合成，反射分量反映了图像的实际内容（如细节、纹理、边缘等），随图像细节不同在空间上做快速变化，其频谱落在空间高频区域。而照明分量在空间上通常均具有缓慢变化的性质，其频谱落在空间低频区域。这样可以通过傅里叶变换将两者分开，进一步进行同态滤波。

原理如图 6.16 所示。其基本思想是按照高通滤波器设计，压缩低频分量，提升高频分量。其假设也是合理的：照明函数虽然频率变化缓慢，但幅度变化大，数字化占用很多比特数，所以要压缩；反射函数描述的是人们感兴趣的景物区频率变化较快，但灰度变化很小、层次不清、细节不明，应该扩展。

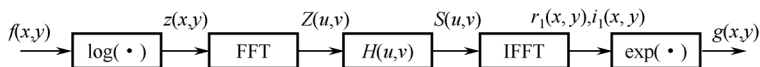


图 6.16 同态增晰原理

6.3.3 增晰算法

增晰算法的步骤如下：

- (1) 对照明函数和反射函数在空间域采用对数运算进行分离

$$z(x, y) = \ln f(x, y) = \ln i(x, y) + \ln r(x, y) \quad (6.22)$$

(2) 通过傅里叶变换将照明函数和反射函数在频率域分开

$$\text{FFT}[z(x, y)] = \text{FFT}[\ln i(x, y)] + \text{FFT}[\ln r(x, y)] \quad (6.23a)$$

设 $Z(u, v) = \text{FFT}[z(x, y)]$, $I_1(u, v) = \text{FFT}[\ln i(x, y)]$, $R_1(u, v) = \text{FFT}[\ln r(x, y)]$, 则上式简写为

$$Z(u, v) = I_1(u, v) + R_1(u, v) \quad (6.23b)$$

(3) 选择同态滤波特性 $H(u, v)$ 。如上所述的原则是：衰减 $I_1(u, v)$ 以压缩 $i(x, y)$ 分量的变化范围，提升 $R_1(u, v)$ 以增强 $r(x, y)$ 细节分量的对比度。同态滤波的输出为

$$S(u, v) = H(u, v)I_1(u, v) + H(u, v)R_1(u, v) \quad (6.24)$$

(4) 通过傅里叶逆变换将滤波后的照明函数和反射函数在空间域分开

$$i_1(x, y) = \text{IFFT}[H(u, v)I_1(u, v)] \quad (6.25a)$$

$$r_1(x, y) = \text{IFFT}[H(u, v)R_1(u, v)] \quad (6.25b)$$

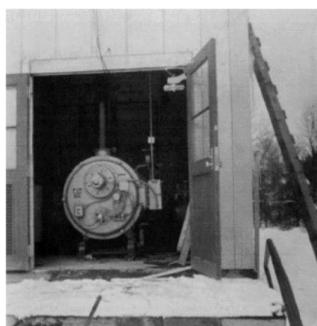
(5) 对照明函数和反射函数在空间域采用指数运算进行还原，将两个新的函数相乘得到增晰的图像

$$i_0(x, y) = \exp[i_1(x, y)] \quad (6.26a)$$

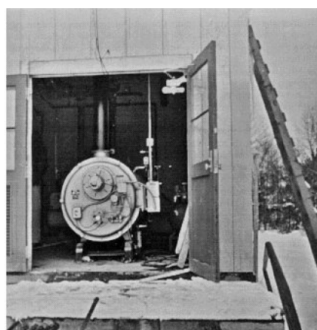
$$r_0(x, y) = \exp[r_1(x, y)] \quad (6.26b)$$

$$g(x, y) = i_0(x, y)r_0(x, y) \quad (6.26c)$$

图 6.17 (a) 所示同态滤波实例是 Thomas Stockham 等人提出同态滤波概念所采用的验证图像，从图 6.17 (b) 可见增晰后的效果是很明显的。



(a) 原始图像



(b) 同态增晰后的图像

图 6.17 同态滤波实例

6.4 平滑

区域增强算法包括平滑 (smoothing) 算法和锐化算法。从频率域看，低通滤波可以对图像进行平滑去噪处理，而高通滤波可对图像进行边缘锐化处理。本节将介绍邻域平均法、中值滤波和边界保持类滤波等。

6.4.1 图像噪声

数字图像往往要经过采集、处理、存储、传输等一系列加工变换，而由电气系统和外界引入的图像噪声也将在这些过程中随之引入，可能严重影响图像的质量。这些过程将使得图像噪声的精确分析

变得十分复杂。图像噪声消除或减低在航空航天、生物医学等领域的图像预处理中的地位显得十分重要,如遥感图像在去噪、去模糊等处理中遗留或产生的影像辐射噪声将影响判读解译的精度,低强度X射线影像系统中的降噪对疾病诊断具有重要意义。

1. 图像噪声的分类

(1) 按其产生的原因,可以分为外部噪声和内部噪声。外部噪声是系统外部干扰以电磁波或经电源串进系统内部而引起的噪声,如电气设备、天体放电现象等引起的噪声。内部噪声是由光和电的基本性质所引起的噪声,如电子器材本身引起的噪声、系统内部设备电路所引起的噪声等。

(2) 按统计特性是否随时间变化,可以分为平稳噪声和非平稳噪声。

(3) 按噪声幅度随时间分布形状来定义,有高斯噪声、瑞利噪声、泊松噪声等。

(4) 按噪声频谱形状来分类,有白噪声、 $1/f$ 噪声、三角噪声等。

(5) 按噪声和信号之间的关系,可分为加性噪声和乘性噪声。

① 加性噪声:假定信号为 f , 噪声为 n , 每一个像素的噪声不管输入信号大小,总是分别叠加在信号上。受噪声污染的信号呈现 $f+n$ 形式。热噪声、散弹噪声、量化噪声等属于加性噪声。

② 乘性噪声:这类噪声受到图像信息的调制,叠加波形为 $f+fn$ 形式。如扫描光栅、胶片颗粒噪声等。在传输过程中,乘性噪声一般由信道不理想引起,而乘性随机性被视为由系统的时变性或者非线性造成。为分析方便起见,一般假定信号和噪声是互相统计独立的。

2. MATLAB 为图像加噪声的函数

MATLAB 中为图像加噪声的函数形式为

$$J = \text{imnoise}(I, \text{type}, \text{parameters})$$

其中, I 为原图像的灰度矩阵, J 为加噪声后图像的灰度矩阵。 type 为噪声种类, parameters 是允许修改的参数,可以默认。 type 可以有 5 种,分别为 'gaussian' (高斯白噪声)、'localvar' (与图像灰度值有关的零均值高斯白噪声)、'poisson' (泊松噪声)、'salt & pepper' (椒盐噪声,即黑白点噪声)和 'speckle' (斑点噪声)。

6.4.2 邻域平均法

1. 空间域分析

大部分噪声可以视为随机信号,它们对图像的影响可以看成是孤立的。对于某一像素而言,如果它与周围像素点相比,有明显的不同,就可以认为该点被噪声感染了。基于这样的分析,我们可以用邻域平均的方法,来判断每一点是否含有噪声,并用适当的方法消除所发现的噪声。图像 $f(x,y)$ 邻域 S 的平均 $g(x,y)$ 可以表示为

$$g(x,y) = \frac{1}{M} \sum_{(x,y) \in S} f(x,y) \quad (6.27)$$

式中, M 为邻域 S 中像素的个数。由此式可知,平均是以图像的模糊为代价来换取噪声的降低的。邻域越大,降噪效果越好。可以证明,平均后噪声的标准差降为原来的 $1/\sqrt{M}$ 。

设当前待处理像素为 $f(x,y)$, 处理后的图像为 $g(x,y)$ 。一个大小为 3×3 的处理模板 (template) 或

掩模 (mask) 为 $H_0 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, 平滑处理过程如图 6.18 所示。

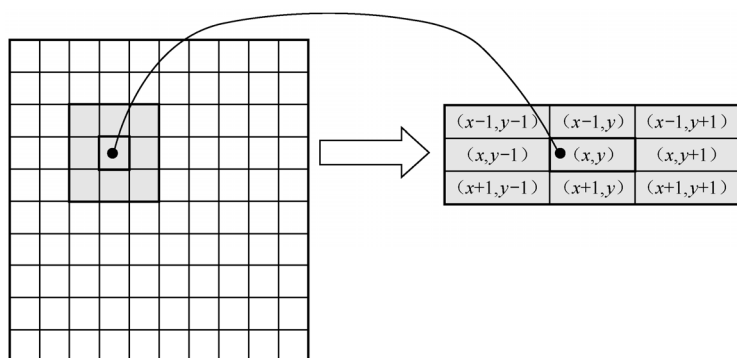


图 6.18 3×3 模板平滑处理示意图

平滑处理过程可描述为

$$g(x,y) = \begin{cases} \frac{1}{M} \sum_{(x,y) \in S} f(x,y), & \left| f(x,y) - \frac{1}{M} \sum_{(x,y) \in S} f(x,y) \right| > T \\ f(x,y), & \text{其他} \end{cases} \quad (6.28)$$

式中, T 为非负阈值, 它可以根据对误差容许的程度, 选为图像灰度均方差 σ_f 的若干倍, 或者通过实验得到。

这种邻域平均的方法也可以用另一种形式来表示, 把平均处理视为图像通过一个低通空间滤波器后的结果。由于噪声多集中在高频段, 而图像的主要能量集中在低频段, 因此采用衰减高频的方法可以有效平滑噪声。设该滤波器的冲激响应为 $H(r,s)$, 于是滤波器输出的结果 $g(x,y)$ 可以表示成卷积 (运算符为 $*$) 的形式, 即

$$g(x,y) = H(x,y) * f(x,y) = \sum_{r=-k}^k \sum_{s=-l}^l f(x-r,y-s) H(r,s) \quad (6.29)$$

式中, $x, y = 0, 1, 2, \dots, N-1$ 。 k 和 l 决定了所选邻域的大小 $(2k+1) \times (2l+1)$ 。 $h(r,s)$ 也可以视为模板 $H_i (i=0, 1, 2, \dots)$ 的元素。除了前面提到的 H_0 平滑模板外, 常用的 3×3 模板如下所示

$$H_1 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad H_2 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad H_3 = \frac{1}{8} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad H_4 = \frac{1}{2} \begin{bmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 1 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{bmatrix}$$

式中, H_2 模板为高斯模板。

MATLAB 采用 `filter2` 函数实现上述的卷积功能, 函数为二维线性数字滤波。主要形式如下

$$G = \text{filter2}(H,F)$$

上述语句中, 使用矩阵 H 中的二维滤波器对输入图像 F 进行 FIR 滤波, 得到输出图像 G 。

【例 6.5】 分别采用上述前 4 种模板对受到椒盐噪声污染的 `eight.tif` 图像进行平滑处理。

【解】 程序如下:

```
I = imread('eight.tif'); %读入原始图像
imshow(I, []);
f = imnoise(I, 'salt & pepper', 0.04); %对图像加椒盐噪声, 噪声强度为 0.04
figure, imshow(f);
h0 = 1/9.*[1 1 1; 1 1 1; 1 1 1]; %定义平滑模板
h1 = [0.1 0.1 0.1; 0.1 0.2 0.1; 0.1 0.1 0.1];
h2 = 1/16.*[1 2 1; 2 4 2; 1 2 1]; %高斯模板
h3 = 1/8.*[1 1 1; 1 0 1; 1 1 1];
```



```

g0 = filter2(h0,f);           %用模板进行滤波处理
g1 = filter2(h1,f);
g2 = filter2(h2,f);
g3 = filter2(h3,f);
figure,imshow(g0,[]);        %显示平滑处理结果
figure,imshow(g1,[]);
figure,imshow(g2,[]);
figure,imshow(g3,[]);

```

图 6.19 显示了原始图像、有噪声图像和 4 种模板处理后的图像。从仿真的效果可以看出,基本的平均模板 H_0 虽然有效消除了目标(硬币)上的椒盐噪声,但也明显损伤了目标的细节。模板 H_1 、 H_2 和 H_3 在兼顾消噪和保持目标的细节的效果方面有不同程度的改善。

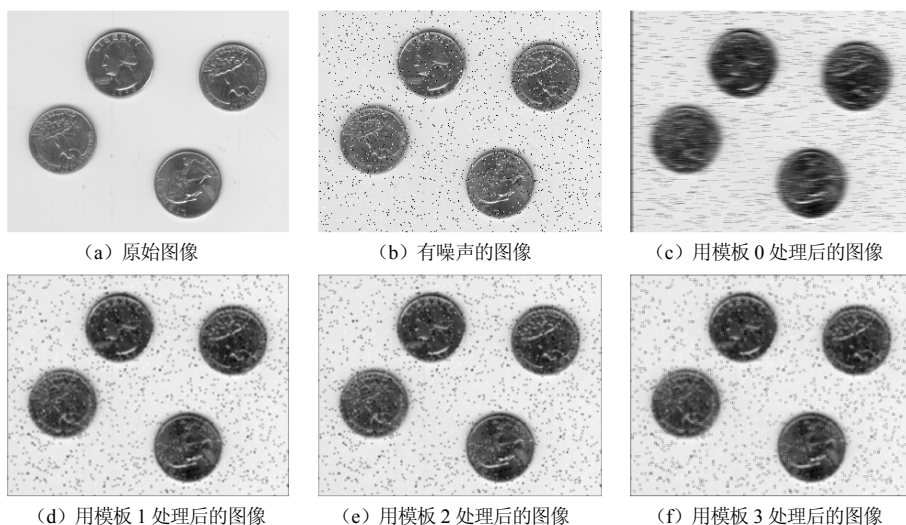


图 6.19 平滑处理的实例

2. 频率域分析

对式 (6.29) 进行二维 DFT, 则将空间域的卷积关系转化为频率域的乘法关系

$$G(u,v) = H(u,v) \cdot F(u,v) \quad (6.30)$$

式中, $G(u,v) = \text{DFT}[g(u,v)]$, $F(u,v) = \text{DFT}[f(u,v)]$, 而 $H(u,v) = \text{DFT}[h(u,v)]$ 为低通滤波器。由于图像的细节也趋向于高频段, 所以选择低通滤波器的截止频率时要特别小心, 兼顾解决降噪和保持图像细节的矛盾。低通滤波器平滑的原理框图如图 6.20 所示。频率域增强的一个示例如图 6.21 所示。

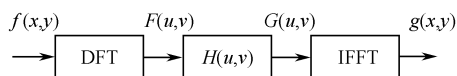


图 6.20 频率域平均去噪原理框图

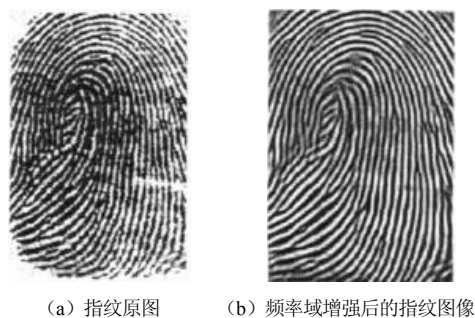


图 6.21 指纹图像的频率域增强

6.4.3 中值滤波

1. 滤波原理

在邻域平均法中，为了抑制噪声，选用了低通滤波器。但是图像的边界通常含有大量的高频信息，所以在去噪的同时，平滑操作也使边界变模糊了。我们希望找到一种在滤除噪声的同时还能保留边缘信息的方法。中值滤波便属于这一类的非线性增强方法，对消除孤立点和线段脉冲等干扰及图像扫描噪声最为有效，但对于消除高斯噪声的影响效果不佳。中值滤波的原理如下。

选一个含有奇数点的滑动窗口 W ，将这个窗口在图像上扫描，把该窗口中所含的像素点按灰度级的升（或降）序排列，取位于中间的灰度值（Med）来代替该点的灰度值。即

$$g(m,n) = \text{Med}\{f(m-k,n-l), (k,l) \in W\} \quad (6.31)$$

如图 6.22 所示，选择滤波用的一维线状窗口 W ，如图 6.22 (a) 所示，待处理像素的灰度取这个模板中灰度的中值，滤波过程如图 6.22 (b) 所示。

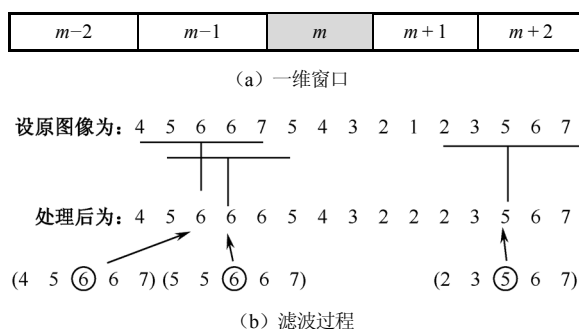


图 6.22 一维窗口及滑动滤波过程

二维中值滤波的窗口形状和尺寸对滤波性能的影响较大，以适应不同的图像内容 and 应用要求。常用的二维窗口有线状、方形、十字形、圆形和环形等。窗口尺寸一般先取 3×3 ，再取 5×5 ，逐步增大，直至达到较满意的滤波效果。对于一些细节较多的复杂图像，还可以多次使用不同的中值滤波进行复合处理，然后通过适当的方式综合所得的结果作为输出，这样可以获得更好的平滑和保护边缘的效果。

2. MATLAB 的二维中值滤波函数

程序中使用到的二维中值滤波器函数 MEDFILT2 的主要形式为

$$J = \text{medfilt2}(I) \text{ 或 } I = \text{medfilt2}(A, [m \ n])$$

前一语句用默认的 3×3 滤波窗口对灰度图像 I 进行二维中值滤波，而后一语句用指定大小为 $m \times n$ 的窗口对图像 I 进行二维中值滤波。

【例 6.6】 选用 3×3 的窗口对椒盐噪声进行中值滤波。

【解】 程序如下：

```
I = imread('eight.tif');
imshow(I);
J = imnoise(I, 'salt & pepper', 0.04);
figure, imshow(J);
K = medfilt2(J);           %二维中值滤波
figure, imshow(K);
```

图 6.23 显示了中值滤波的结果。与上例的平滑滤波相比较,可见此时采用中值滤波效果是十分明显的。

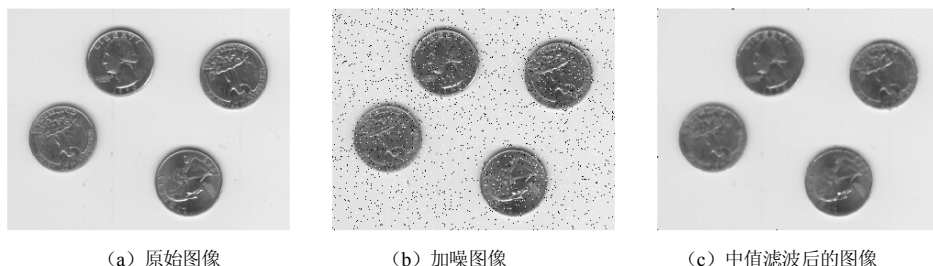


图 6.23 中值滤波

以 SBS 改性沥青材料显微图像滤波为例。在加入 SBS 改性剂和稳定剂后,每隔 15 分钟需要用显微镜对材料进行采样,可观察到颗粒的形状会发生改变,颗粒会慢慢扩散,所以在对具体一张显微图像进行分析时,可以把其中微小的颗粒看作噪声滤波去除。图 6.24 所示为采用均值滤波和中值滤波,且每种滤波方式都可以采用不同的模板进行滤波。由效果图可见,中值滤波对噪声处理效果较好,去除了图中微小的可以看作为杂质的颗粒,并且利用 5×5 模板比利用 3×3 模板去除的杂质点多,但是图像会有些模糊;而均值滤波后图像变得模糊,且利用 5×5 模板的效果比利用 3×3 模板的效果更加模糊,虽在一定程度上抑制了噪声,但由于平均化使得图像明显模糊。

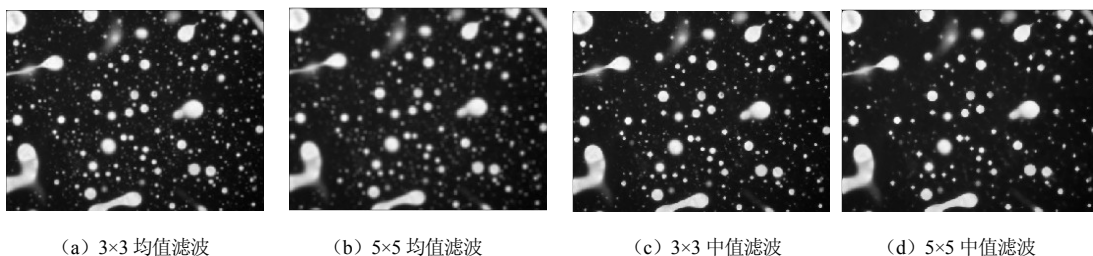


图6.24 显微图像滤波

6.4.4 边界保持类滤波

1. K 近邻均值滤波器 (KNNF)

该算法的思想是:在 $m \times m$ 的窗口中,属于同一集合类的像素,它们的灰度值将高度相关。基于此思想,被处理的像素(对应于窗口中心的像素)可以用窗口内与中心像素灰度最接近的 K 个邻近像素的平均灰度来代替。

处理步骤如下:

- (1) 做一个 $m \times m$ 的作用模板。
- (2) 在其中选择 K 个与待处理像素的灰度差最小的像素。
- (3) 用这 K 个像素的灰度均值替换掉原来的值。

图 6.25 是模板为 3×3 、 $K = 3$ 的 K 近邻均值滤波器的例子,其中求出的均值为 2.67,经过取整处理后得到像素的值为 3。

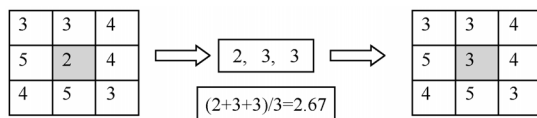


图 6.25 K 近邻均值滤波器

2. K 近邻中值滤波器 (KNNMF)

在 K 近邻均值滤波器中, 不选 K 个邻近像素的平均灰度来代替, 而选 K 个邻近像素的中值灰度来代替, 则这个滤波器就变成了 K 近邻中值滤波器。

3. 最小均方差滤波器

对图像上待处理的当前像素, 选其 5×5 邻域, 在此邻域中采用图 6.26 所示的模板, 计算各个模板的均值 $\bar{f} = \frac{1}{N} \sum_{(x,y) \in S} f(x,y)$ (其中 S 是对应的模板, N 是模板中像素的数量) 和方差

$$\sigma^2 = \frac{1}{N} \sum_{(x,y) \in S} [f(x,y) - \bar{f}]^2$$

按方差排序, 最小方差所对应的模板的灰度均值就是当前像素的输出值。

该方法以方差作为各个邻域灰度均匀性的测度。若邻域含有尖锐的边缘, 它的灰度方差必定很大, 而不含边缘或灰度均匀的邻域, 它的方差就很小, 从而最小方差所对应的邻域就是灰度最均匀的邻域。因此, 通过这样的平滑既可以消除噪声, 又能够不破坏邻域边界的细节。

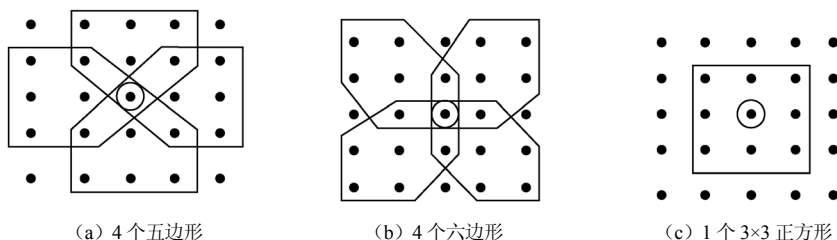


图 6.26 最小均方差滤波器模板

6.5 锐化

图像锐化处理 (sharpening) 的目的是加强图像中景物的边缘和轮廓, 使模糊图像变得更清晰。图像模糊的实质是由于图像受到平均或积分运算, 因此对其采用逆运算, 如对连续图像微分或对离散图像差分运算, 就可以使模糊图像的质量得到改善。从频率域角度看, 图像的模糊是其高频分量受到衰减, 因而采用合适的高通滤波器可以使图像增晰。值得注意的是, 待锐化的图像要有足够的信噪比, 否则会使噪声得到比原图像更强的增强, 信噪比更加恶化。

6.5.1 空间域差分法

图像的边缘和轮廓一般位于中灰度突变的区域, 因而可以用灰度的差分提取边缘和轮廓并进行增强。

1. 梯度锐化法

梯度的方向在二元连续函数 $f(x, y)$ 于坐标点 (x, y) 处的最大变化方向上, 幅度是其最大变化率方向上的单位距离所增加的量。梯度定义为

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (6.32)$$

其相角 Ang 和幅度 Mag 分别由下式给出

$$\text{Ang}(\nabla f) = \arctan \left(\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}} \right) \quad (6.33a)$$

$$\text{Mag}(\nabla f) = |\nabla f| = [G_x^2 + G_y^2]^{\frac{1}{2}} = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{\frac{1}{2}} \quad (6.33b)$$

尽管梯度向量的偏导数分量本身是线性算子，具有移位不变的性质，但并非是各向同性的，为了把图像各方向模糊的边缘和轮廓线变清晰，我们还希望对图像的某种导数运算具有各向同性。梯度幅度呈现非线性性质，但具有各向同性或旋转不变性。设 (x, y) 是图像旋转前的坐标， (x', y') 是图像旋转后的坐标，可以证明下述的梯度幅度旋转不变性

$$\left(\frac{\partial f}{\partial x'} \right)^2 + \left(\frac{\partial f}{\partial y'} \right)^2 = \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \quad (6.34)$$

为了降低运算量且保持灰度的相对变化，常用绝对值或最大值运算代替平方与平方根运算来近似求梯度的幅度

$$|\nabla f| \approx |G_x| + |G_y| \quad (6.35a)$$

$$|\nabla f| \approx \text{Max}(|G_x|, |G_y|) \quad (6.35b)$$

对于数字图像，用差分代替微分，沿 x 和 y 方向（分别替换成离散坐标 i 和 j ）的一阶差分分别表示为

$$G_x = f(i, j+1) - f(i, j) \quad (6.36a)$$

和

$$G_y = f(i+1, j) - f(i, j) \quad (6.36b)$$

沿 x 和 y 方向的一阶差分的示意图见图 6.27。

所以数字图像的差分运算可取

$$|\nabla f(i, j)| \approx |G_x| + |G_y| = |f(i, j) - f(i, j+1)| + |f(i+1, j) - f(i, j)| \quad (6.37a)$$

或

$$|\nabla f(i, j)| \approx \text{Max}(|G_x|, |G_y|) = \text{Max}(|f(i, j+1) - f(i, j)|, |f(i+1, j) - f(i, j)|) \quad (6.37b)$$

梯度也可以采用罗伯茨（Roberts）交叉梯度表示，如图 6.28 所示，其表达式如下

$$|\nabla f(i, j)| \approx |f(i+1, j+1) - f(i, j)| + |f(i+1, j) - f(i, j+1)| \quad (6.38a)$$

或

$$|\nabla f(i, j)| \approx \text{Max}(|f(i+1, j+1) - f(i, j)|, |f(i+1, j) - f(i, j+1)|) \quad (6.38b)$$

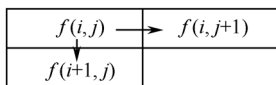


图 6.27 沿 x 和 y 方向的一阶差分

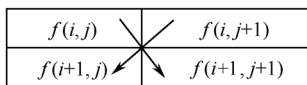


图 6.28 罗伯茨差分

所有梯度值都和相邻像素之间的灰度差分成比例，因而在灰度变化比较大的边界轮廓点处有较大的梯度值，而在灰度变化比较平缓的区域，相应的梯度值也较小。因此我们可以利用它来增强图像中景物的边界，达到锐化的目的。采用梯度进行图像锐化的方法有以下几种。

(1) 输出图像 $g(i, j)$ 的各点等于该点处的梯度。即

$$g(i, j) = |\nabla f(i, j)| \quad (6.39)$$

这种方法的缺点是，输出的图像在灰度变化比较小的区域， $g(i,j)$ 很小，显示的是一片黑色。

(2) 对梯度值超过某阈值 T 的像素选用梯度值，而小于该阈值时保持原图像的像素值，即

$$g(i,j) = \begin{cases} |\nabla f(i,j)|, & |\nabla f(i,j)| \geq T \\ f(i,j), & \text{其他} \end{cases} \quad (6.40)$$

适当地选取 T ，可以有效地增强边界而不影响比较平滑的背景。

(3) 对梯度值超过某阈值 T 的像素选用固定灰度 L_G 代替，而小于该阈值时仍选用原图像的像素点值，即

$$g(i,j) = \begin{cases} L_G, & |\nabla f(i,j)| \geq T \\ f(i,j), & \text{其他} \end{cases} \quad (6.41)$$

这种方法可以使边界清晰，同时又不损害灰度变化比较平缓区域的图像特性。

(4) 将梯度值超过某阈值 T 的像素选用梯度值，而小于该阈值时选用固定的灰度 L_B ，即

$$g(i,j) = \begin{cases} |\nabla f(i,j)|, & |\nabla f(i,j)| \geq T \\ L_B, & \text{其他} \end{cases} \quad (6.42)$$

这种方法将背景用一个固定的灰度级 L_B 来表示，便于研究边缘灰度的变化。

(5) 将梯度值超过某阈值 T 的像素选用固定灰度 L_G ，而小于该阈值时选用固定的灰度 L_B ，即

$$g(i,j) = \begin{cases} L_G, & |\nabla f(i,j)| \geq T \\ L_B, & \text{其他} \end{cases} \quad (6.43)$$

这种方法生成的是二值图，根据阈值将图像分成边缘和背景，便于研究边缘所在的位置。

在实际应用中，常常采用小型模板，然后利用卷积运算来近似梯度， G_x 和 G_y 各自使用一个模板。对模板有一些基本要求：模板中心的系数为正，其余相邻系数为负，且所有的系数之和为零。人们已经提出了许多不同大小、不同系数的模板，最简单的是上述的 Roberts 算子，其 G_x 和 G_y 模板如式 (6.44) 所示。

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad G_y = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (6.44)$$

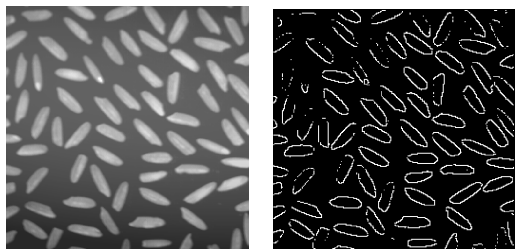
【例 6.7】 利用罗伯茨梯度对 rice.tif 图像进行锐化处理。

【解】 程序如下：

```
I = imread('rice.tif');
imshow(I);
BW = edge(I,'roberts',0.1); %对输入图像求罗伯茨梯度
figure, imshow(BW);
```

程序中使用的 edge 函数用于灰度图像的边缘检测，主要形式为

$$BW = \text{edge}(I, \text{method}, \text{thresh})$$



(a) 原图像

(b) 锐化结果图

图 6.29 罗伯茨梯度的锐化

edge 函数对灰度图像 I 进行边缘检测，检测的方法由 method 决定，检测输出的图像是二值图像，阈值由 thresh 指定。属于上述的方法 (5)。method 除了选择 'roberts' 外，还支持其他 5 种不同形式，即 'sobel'、'prewitt'、'log'、'zerocross' 和 'canny'。本函数的具体应用将在第 8 章中详述。

图 6.29 是输入的 “rice.tif” 图像和处理的结果图。

表 6.2 中给出了常用的梯度算子的对比。

表 6.2 常用的梯度算子

算 子	G_x	G_y	特 点
Roberts	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	定位准确, 但对噪声敏感
Prewitt	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	用像素点上、下、左、右邻点的灰度差, 在边缘处达到极大值检测边缘, 去掉部分伪边缘, 对噪声具有平滑作用。与 Roberts 相比, 减低了对噪声的敏感程度
Sobel	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	与上述算子相比, 对于像素的位置的影响做了加权, 因此效果更好。提取的图像轮廓有时并不能令人满意
Isotropic Sobel	$\begin{bmatrix} -1 & -\sqrt{2} & -1 \\ 0 & 0 & 0 \\ 1 & \sqrt{2} & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix}$	又称为各向同性 Sobel 算子。和 Sobel 算子相比, 它的位置加权系数更为准确, 在检测不同方向的边沿时梯度的幅度一致
Kirsch	$\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}$	$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}$	共有 8 个方向算子, 这里列出的是水平和垂直两个方向的算子, 有较好的抑制噪声的作用

2. 拉普拉斯算子

拉普拉斯算子 (Laplacian) 是具有各向同性的二阶微分算子。一个连续的二元函数 $f(x, y)$, 其拉普拉斯运算定义为

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (6.45)$$

对于由扩散现象引起的图像模糊过程, 如胶片颗粒化学扩散、光点散射等, 可以模型化为下述方程

$$g = f - k\tau \cdot \nabla^2 f \quad (6.46)$$

式中, f 和 g 分别为锐化前后的图像; $k\tau$ 为常数。可见, 锐化后的图像是有模糊图像减去拉普拉斯算子乘以 $k\tau$ 得到的。 $k\tau$ 反映了锐化的程度, 若太大, 会使图像中的轮廓产生过冲; 而若过小, 锐化效果会不明显。

对于数字图像, 拉普拉斯算子锐化过程可以简化为

$$g(i, j) = 4f(i, j) - f(i+1, j) - f(i-1, j) - f(i, j+1) - f(i, j-1) \quad (6.47)$$

上式也可以表示为卷积的形式, 即

$$g(i, j) = \sum_{r=-k}^k \sum_{s=-l}^l f(i-r, j-s) H(r, s) \quad (6.48)$$

式中, $i, j = 0, 1, 2, \dots, N-1$, $k=1$, $l=1$, $H(r, s)$ 如下:

$$H_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (6.49)$$

在图像处理过程中, 函数的拉普拉斯算子也是借助模板来实现的。常用的模板有

$$G_x = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \quad G_y = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (6.50)$$

图 6.30 给出了拉普拉斯算子图像锐化的效果比较。



(a) 原 Cameraman 图像

(b) 锐化后的图像

图 6.30 拉普拉斯算子图像锐化实例

IPT 提供了对任意类型数组或多维图像进行滤波的 `imfilter` 函数，其调用的主要形式为

$$G = \text{imfilter}(F, H)$$

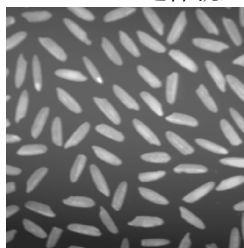
上述的 `imfilter` 函数将灰度图像 F 通过滤波器 H 进行滤波，结果放在 G 中。 H 也可以调用 `fspecial` 函数，调用的基本形式为 $H = \text{fspecial}(\text{TYPE})$ ，其中 TYPE 是滤波器的类型，可选值有：'average'、'disk'、'gaussian'、'laplacian'、'log'、'motion'、'prewitt'、'sobel'和'unsharp'。

【例 6.8】应用拉普拉斯算子对 `rice.tif` 图像进行锐化处理。

【解】程序如下：

```
I = imread('rice.tif');
imshow(I);
H = [0 -1 0; -1 4 -1; 0 -1 0];
J = imfilter(I, H);
figure, imshow(J);
figure, imhist(J);
K = imadjust(J, [0.0 0.2], []);
figure, imhist(K);
figure, imshow(K);
```

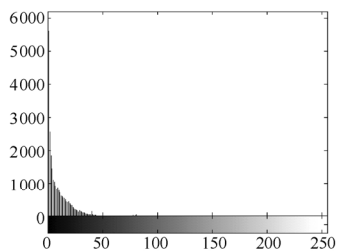
锐化处理的结果见图 6.31。由于锐化后的图像（见图 6.31 (b)）亮度偏低（见图 6.31 (c)），因而进行对比度扩展（见图 6.31 (d)），这样就生成了边缘清晰的图像。



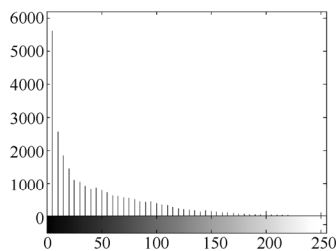
(a) 原始图像



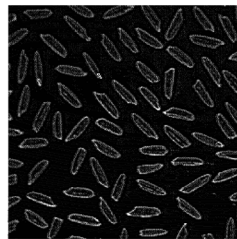
(b) 拉普拉斯锐化



(c) 拉普拉斯锐化后图像的直方图



(d) 对锐化后图像的对比度扩展



(e) 对比度扩展后的图像

图 6.31 拉普拉斯算子的锐化

以 SBS 改型沥青材料显微图像为例。用图像锐化的方法突出颗粒轮廓部分，以此来锐化图像中由邻域平均造成的图像模糊。采用的无方向微分锐化算子，如图6.32。四个算子比较可见 LoG 算子（第8章将详述）对该粒子图片的处理效果较清晰。

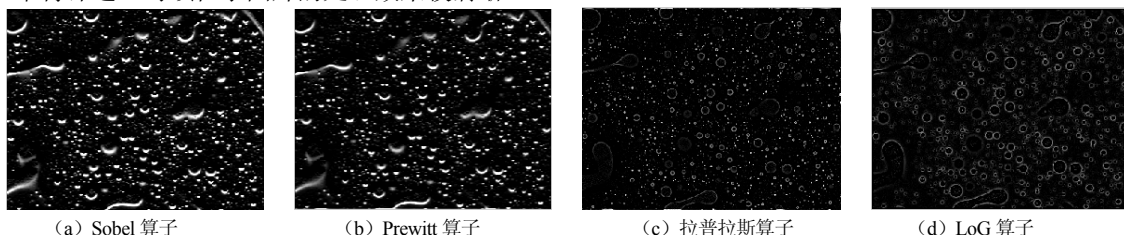


图6.32 显微图像锐化

6.5.2 频率域高通滤波法

如上所述，图像模糊的本质是图像的高频分量受到衰减，采用合适的高通滤波器提升高频分量将会使模糊得到相应的补偿。只要适当地选择滤波因子 $H(r, s)$ ，就可以组成不同性质的高通滤波器，从而使图像达到期望中的增强效果。常用的高通模板有

$$H_2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (6.51)$$

6.6 实验：图像增强

1. 实验目的

- (1) 熟悉并学会使用 MATLAB 中图像增强的相关函数及 Photoshop 操作。
- (2) 掌握图像灰度修正、平滑去噪、锐化加强边缘和轮廓的方法，并编程实现。

2. 实验主要仪器设备

- (1) 台式机或笔记本电脑。
- (2) MATLAB 软件（含图像处理工具箱）、Photoshop。注意：由于软件版本的缘故，软件的界面可能有所差异，读者可以根据实际安装的软件选择相关的命令。
- (3) 典型的灰度、彩色图像文件。

3. 实验原理

数码相机的曝光量指到达 DC 感光器件上的光线总量，用曝光值 (EV) 表示。图像的过度曝光、曝光不足时，用曝光补偿调节曝光量，这种功能可修正自动曝光设置值为上升或下降几级。例如，某些 DC 的 EV 调整范围为 +3~0~-3。尝试对同一景象进行正确曝光、过度曝光和曝光不足三种情况成像情况。

- (1) 将一幅图像视为一个二维矩阵，用 MATLAB 进行图像增强。
- (2) 利用 MATLAB 图像处理工具箱中的函数 `imread`（读）、`imshow`（显示）、`imnoise`（加噪）、`filter2`（滤波）对图像进行去噪处理。
- (3) 图像灰度修正：灰度变换。对不满意的图像通过线性或非线性灰度映射关系进行变换，其效果可以得到明显提高。通过分析，会发现变换前后图像的直方图也发生相应的变化。

(4) 图像平滑方法：邻域平均、中值滤波。分析图像降质的性质，区分平稳型还是非平稳型、加性还是乘性等，采用合适的去噪方法，可以去除或降低噪声对图像的影响。从频率域看，平均操作在减低噪声的同时衰减了图像的高频分量，会影响图像细节的重现。中值滤波对某些信号具有不变性，适用于消除图像中的突发干扰，但如果图像含有丰富的细节，则不宜使用。

(5) 图像锐化方法：人眼对目标的边缘和轮廓较为敏感，对图像进行锐化，有助于突出图像的这些特征。从频率域看，锐化提升了图像的高频分量。

4. 实验内容

(1) MATLAB 图像增强：① 图像灰度修正；② 图像平滑方法；③ 图像锐化方法。

(2) Photoshop 图像增强：① 直方图的计算；② 图像增强；③ 颜色的调整；④ 选择区域工具；⑤ 图层的操作。

5. 实验步骤

(1) MATLAB 图像增强。

① 图像灰度修正。测试图像为 `pout.tif`、`tire.tif`。读入一幅灰度级分布不协调的图像，分析其直方图。根据直方图，设计灰度变换表达式，或调用 `imadjust` 函数。调整变换表达式的参数，直到显示图像的灰度级分布较均衡为止。

② 不均匀光照的校正。测试图像为 `pout.tif`。采用分块处理函数 `blkproc` 和图像相减函数 `imsubtract` 校正图 6.6 存在的不均匀光照现象。

③ 三段线性变换增强。测试图像为 `couple.tif`。选择合适的转折点，编程对图 6.10 进行三段线性变换增强。

④ 图像平滑方法。测试图像为 `eight.tif`。对有噪图像或人为加入噪声的图像进行平滑处理。根据噪声的类型，选择不同的去噪方法，如邻域平均、中值滤波等方法，调用 `filter2`、`medfilt2` 函数，选择不同的滤波模板和参数，观测和分析各种去噪方法对不同噪声图像处理的去噪或降噪效果。

⑤ 图像锐化方法。测试图像为 `rice.tif`、`cameraman.tif`。读入一幅边缘模糊的图像，利用罗伯茨梯度对图像进行 4 种锐化处理，比较各自的效果。

(2) Photoshop 图像增强。选择图 6.28 所示的彩色图像 `Lena` 进行增强，包括分别调整图像的有关参数 [亮度 (Brightness)、对比度 (Contrast)、灰度系数 γ 等]。图 6.33 所示为待增强的图像和相关命令。

6. 实验报告要求

(1) 说明利用 MATLAB 图像处理工具箱实现图像灰度修正、图像平滑、图像锐化的方法。

(2) 列出上述图像处理的程序。

(3) 记录灰度修正、图像平滑、图像锐化的图像。

(4) Photoshop 图像增强的步骤和相应界面。

(5) 回答思考题。

(6) 心得和体会。

7. 预习要求

(1) 了解 MATLAB 图像处理工具箱关于图像增强的有关功能。

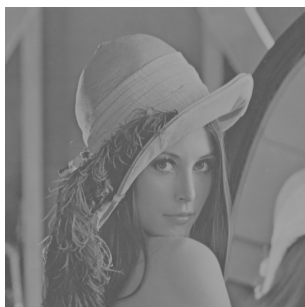
(2) 列出上述图像处理的流程。

(3) Photoshop 图像增强的步骤和相应界面。

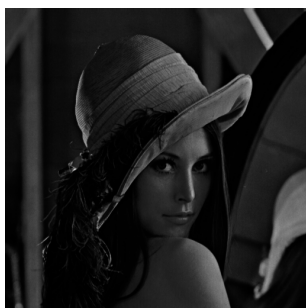
① 直方图的计算。打开需要计算和显示的图像，选择 `Image` (图像) 菜单下的 `Histogram` (直方图) 命令。



(a) 过亮的图像



(b) 过弱的对比度



(c) 过小的灰度系数



(d) 与图像增强相关的命令 (CS 以前版本)

图 6.33 待增强的图像和相关命令

② 图像增强。打开需要增强的图像，选择“图像”菜单下的“调整”子菜单，选择相关的增强命令。

- “亮度/对比度”命令可以独立地调整图像的亮度和对比度。
- “曲线”命令可以调整图像的灰度系数（即 γ 曲线）。
- “自动灰阶”可以实现直方图的均衡化。
- 实验时可以尝试其他命令，观看图像的变化情况，判断是否有增强的效果。

③ 颜色的调整。打开需要进行颜色调整的图像，在 Image（图像）菜单的 Adjust（调整）子菜单下选择相关的颜色调整命令。

- 亮度和对比度的调整与上述的图像增强相同。
- 选择 Variations…（变化……）命令可以进行图像的变色。

8. 思考题

- (1) 如何针对图像过暗、过亮、对比度不足设计灰度变换函数？
- (2) 比较同一种去噪方法对不同噪声处理的效果。
- (3) 讨论用梯度法锐化图像的 4 种不同方法的应用范围。
- (4) 结合本书介绍的图像增强方法，说明 Photoshop 在此方面具有的功能。

本章小结

本章介绍了图像增强的基本方法，有空间域增强和频率域增强两种方法。作为基本的图像处理方法，图像增强具有重要地位。后期的图像处理往往在图像增强后完成。空间域增强直接对像素灰度进行操作，而频率域增强需要对图像进行傅里叶变换等处理过程。

本章在讲解基本概念和方法的同时,给出了许多具体实现的例子,旨在让读者结合增强算法的实现和运行结果来理解相关的方法与适用范围。本章的重点是对各种增强方法的理解和应用,包括图像的灰度变换、图像的去噪方法和图像锐化的各种算法。难点是在图像处理时如何选用增强方法,使得图像得到较好的增强,在这方面对图像的深入分析和细心观察是不可缺少的,需要不断积累和借鉴自己和他人的经验。读者可以尝试完成本章设计的图像增强实验,以加深对相关内容的理解。

思考题与习题

6.1 图像增强的目的是什么?它包含哪些研究内容?与图像复原有何区别?

6.2 试分别给出将灰度范围(10,100)拉伸到(0,150)和(10,200)、压缩到(50,100)和(10,125)、平移到(110,225)的线性变换。

6.3 如图 6.34 所示,已知灰度图像 $f(x,y)$,求经过反转变换后的图像 $g(x,y)$ 。反转变换为 $g = G(f)$ 。

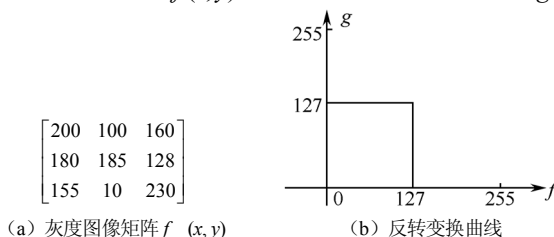


图 6.34 灰度图像的反转

6.4 试求灰度变换方程 $g = G(f)$,使其在范围 $10 \leq f \leq 150$ 内为 $\ln(f)$ 的线性函数。

6.5 设有一幅 64×64 的离散图像,其灰度分成 8 层,灰度 n_k 的值和分布情况如表 6.3 所示。试绘制该图像的直方图,并求经过直方图均衡后的图像的直方图。说明为什么对数字图像进行直方图均衡化后,通常并不能产生完全平坦的直方图。

表 6.3 一幅图像的灰度分布

K	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	560	920	1046	705	356	267	170	72

6.6 试述图像平滑的目的、采用空间域和频率域滤波的原理。

6.7 试述中值滤波的特点。为什么对一些细节多,特别是点、线、尖点细节多的图像不适宜用这种方法?

6.8 如图 6.35 所示,设原图像为 10×10 的点阵,求边界点保持不变、经过 3×3 窗口中值滤波的图像。

```

1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1
1 1 5 5 5 5 5 5 1 1
1 1 5 5 5 5 5 5 1 1
1 1 5 5 8 8 5 5 1 1
1 1 5 5 8 8 5 5 1 1
1 1 5 5 5 5 5 5 1 1
1 1 5 5 5 5 5 5 1 1
1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1

```

图 6.35 受干扰的二维图像

6.9 设图像如图 6.36 所示, 分别求经过邻域平滑和高通算子锐化的结果。其中边缘点保持不变, 邻

域平滑掩码取 3×3 矩阵, 即 $H = \frac{1}{8} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$; 邻域高通算子取 3×3 矩阵, 即 $H = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$ 。

1	1	3	4	5
2	1	4	5	5
2	3	5	4	5
3	2	3	3	2
4	5	4	1	1

图 6.36 一幅 5×5 的灰度图像矩阵

6.10 为什么待锐化的图像需要有足够的信噪比? 选择不同信噪比 ($\text{SNR} = 40 \text{ dB}$ 和 20 dB) 的图像用 Roberts 梯度法实验说明。

6.11 已知一幅如图 6.37 所示的图像。可见原图中左边暗, 右边亮, 中间存在着一条明显的边界。

```

0 0 1 255 254 254 254
1 1 1 254 253 254 254
0 0 0 255 255 253 253
1 1 0 254 254 254 254

```

图 6.37 有垂直边界的一幅图像

试用 Sobel 算子对给定的图像进行模板操作并分析得到的结果图像。

第7章 图像复原

内容提要

本章介绍数字图像复原的基本任务、图像退化的各种原因、图像复原的常用方法和相关的 IPT 函数。通过本章的学习，读者应逐步掌握退化现象模型化和解决方案公式化的技能。

知识要点

- 图像退化的常见原因。
- 图像退化模型。
- 线性代数复原。
- 非线性复原。
- 盲图像复原。
- 几何失真校正。
- 图像修复简介。

教学建议

- 本章教学安排 4 学时，在内容上与“图像增强”有些交叉，但侧重点不同。近年来，图像修复技术成为新的研究和应用技术，图像修复和图像复原在目的和方法上有相似之处，但根据不同的应用需求，修复后的图像可能逼近原始图像，也可能带有主观目的进行修复。
- 重点了解数字图像复原的基本任务、图像退化的各种原因、图像复原的常用方法，能够用 IPT 去模糊函数等资源解决简单的图像退化问题。
- 学习本课程的先修知识包括线性代数（循环矩阵的表示、矩阵的广义逆等）、信号与线性系统，数字信号处理（圆周卷积、离散卷积定理等）、随机过程（平稳随机过程等）、优化理论等。
- 本章涉及的数学知识较多且较难。读者可以根据教学要求、自己的兴趣或实际工作需要，逐步了解和消化相关内容，重点放在解决问题的思路和方法上，力图用工程方法解决实际遇到的图像退化问题。本章的三个实例提供了用 IPT 求解图像复原问题的方法。读者可以尝试提出解决类似问题的方法。
- 注意本章与“图像增强”一章的联系与区别。
- 非线性复原方法、图像修复可作为选讲内容。

7.1 图像退化原因与复原技术分类

图像在形成、传输和记录等过程中, 由于受到多方面的影响, 不可避免地造成图像质量的退化(degradation)。造成图像退化的原因很多, 大致可分为以下几个方面:

- (1) 射线辐射、大气湍流、雾霾天气等造成的照片畸变或降质。
- (2) 模拟图像数字化的过程中, 由于会损失部分细节, 造成图像质量下降。
- (3) 镜头聚焦不准产生的散焦模糊。
- (4) 成像系统中始终存在的噪声干扰。
- (5) 拍摄时, 相机与景物之间的相对运动产生的运动模糊。
- (6) 在图像采集过程中电气设备等可能形成图像的周期噪声。
- (7) 底片感光、图像显示时会造成记录显示失真。

(8) 成像系统的像差、非线性畸变、有限带宽等造成的图像失真。如鱼镜头(fisheye lens)具有焦距短、视场角大的优点, 被广泛应用在全视觉监视、机器人导航等全方位视觉系统中, 但用鱼眼摄像机拍摄的图像会严重的变形失真。

- (9) 携带遥感仪器的飞机或卫星运动的不稳定, 以及地球自转等因素引起的照片几何失真。

图像复原(image restoration)是在研究图像退化原因的基础上, 以退化图像为依据, 根据一定的先验知识, 建立一个退化模型, 然后用相反的运算, 来恢复原始景物图像。图像复原要明确规定质量准则, 即衡量接近原始景物图像的程度。图像复原模型可以用连续数学或离散数学处理。图像复原根据退化的数学模型对退化图像进行处理, 其实现可在空间域卷积或在频率域相乘。

由于引起退化的原因多而且性质不同, 描述图像退化过程所建立的数学模型也是各不相同的。加上用于复原的估计准则不同, 因此图像复原的方法、技术也各不相同。一般的图像处理面对的正问题, 即对输入图像进行加工、处理, 得到所需的输出图像。而图像复原显然是一个信号的求逆问题。逆问题一般要比正问题的求解难得多, 常常得不到唯一解, 甚至无解。为了得到逆问题的有用解, 经常需要一些额外的先验知识以及对解的一些其他附加约束条件。可见图像复原是一个复杂的数学求解过程, 必要时需要采用人机结合的方法进行交互式图像恢复。

由于图像复原的目的是要尽可能恢复被退化图像的本来面目, 所以又叫图像恢复。图像复原与图像增强有类似的地方, 二者的目的都是为了改善图像的质量, 但它们追求的目标不同。图像增强不考虑图像是如何退化的, 而是试图采用各种技术来增强图像的视觉效果。因此, 图像增强可以不考虑增强后的图像是否失真, 只要满足人眼或机器视觉的要求即可。而图像复原需要了解图像退化的机制和过程等先验知识, 据此找出一种相应的逆处理方法, 从而得到恢复的图像。如果图像已退化, 应先做复原处理, 再做增强处理。

图像复原在初级视觉处理中占有极其重要的地位, 在航空航天、国防公安、生物医学、文物修复等领域具有广泛的应用。传统的复原方法是基于平稳图像、线性空间不变的退化系统、图像和噪声统计特性的先验知识已知等条件下讨论的, 而现代的复原方法已经在非平稳图像(如卡尔曼滤波)、非线性方法(如神经网络)、信号与噪声的先验知识未知(如盲图像复原)等前提下开展了卓有成效的工作, 取得了令人鼓舞的成果。

7.1.1 连续图像退化的数学模型

连续图像退化的一般模型如图 7.1 所示。输入图像 $f(x, y)$ 经过一个退化系统或退化算子 $H(x, y)$ 后, 产生的退化图像 $g(x, y)$ 可以表示为

$$g(x, y) = H[f(x, y)] \quad (7.1)$$

如果仅考虑加性噪声的影响, 则退化图像可表示为

$$g(x, y) = H[f(x, y)] + n(x, y) \quad (7.2)$$

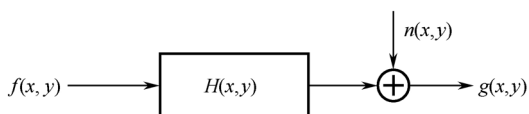


图 7.1 图像退化一般模型

由式 (7.2) 可见, 退化的图像是由成像系统的退化加上额外的系统噪声而形成的。根据此模型可知, 若已知 $H(x, y)$ 和 $n(x, y)$, 图像复原是在退化图像的基础上做逆运算, 得到 $f(x, y)$ 的一个最佳估计 $\hat{f}(x, y)$ 。之所以说是“最佳估计”而非“真实估计”, 是由于

存在以下两个原因可能导致图像复原的病态性。

(1) 进行逆运算时, 最佳估计问题不一定有解。这是由于图像复原中可能遇到最棘手的问题——奇异问题。

(2) 逆问题可能存在多个解。

一幅连续图像 $f(x, y)$ 可以用下式表示

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta \quad (7.3)$$

式中, δ 函数表示空间上点脉冲的冲激函数。

将式 (7.3) 代入式 (7.1) 得

$$g(x, y) = H[f(x, y)] = H \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta \right] \quad (7.4)$$

在退化算子 H 表示线性和空间不变系统的情况下, 输入图像 $f(x, y)$ 经退化后的输出为 $g(x, y)$

$$\begin{aligned} g(x, y) &= H[f(x, y)] = H \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta \right] \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) H[\delta(x - \alpha, y - \beta)] d\alpha d\beta \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta \end{aligned} \quad (7.5)$$

式中, $h(x, y)$ 称为退化系统的冲激响应函数。在图像形成的光学过程中, 冲激为一光点, 而光学部件的响应却是扩散了的模糊光斑, 因而又将 $h(x, y)$ 称为退化系统的点扩展函数 (Point-Spread Function, PSF)。通过冲激光点与系统响应的关系, 我们可以采用实验估计的方法获取成像系统的 PSF。而在有些退化的场合, 可以采用图像观察估计、数学建模估计等方法获取 PSF。

可见, 退化系统的输出就是输入图像 $f(x, y)$ 与点扩展函数 $h(x, y)$ 的卷积, 考虑到噪声的影响, 即

$$\begin{aligned} g(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta + n(x, y) \\ &= f(x, y) * h(x, y) + n(x, y) \end{aligned} \quad (7.6)$$

对上述方程取傅里叶变换, 则式 (7.6) 在频率域上可以写成

$$G(u, v) = F(u, v)H(u, v) + N(u, v) \quad (7.7)$$

式中, $G(u, v)$ 、 $F(u, v)$ 、 $N(u, v)$ 分别是 $g(x, y)$ 、 $f(x, y)$ 、 $n(x, y)$ 的傅里叶变换; $H(u, v)$ 是 $h(x, y)$ 的傅里叶变换, 为系统的传递函数。例如, 一种基于大气湍流物理特性估计的传输函数模型为 $H(u, v) = \exp[-k(u^2 + v^2)^{5/6}]$, 式中, k 是湍流强度有关的常数, 中等湍流的数量级为 10^{-3} 。

7.1.2 离散图像退化的数学模型

数字图像处理系统处理的图像是离散图像,因此我们对式(7.6)的离散形式更感兴趣。在数字信号处理课程中,我们已经讨论了类似系统的一维情况的表示,这里着重将一维离散退化模型推广到二维。

设输入的数字图像 $f(x,y)$ 大小为 $A \times B$,点扩展函数 $h(x,y)$ 被均匀采样为 $C \times D$ 大小。为避免交叠误差,采用添零延拓的方法,将它们扩展成 $M = A + C - 1$ 和 $N = B + D - 1$ 个元素的周期函数

$$f_e(x,y) = \begin{cases} f(x,y), & 0 \leq x \leq A-1 \text{ 且 } 0 \leq y \leq B-1 \\ 0, & \text{其他} \end{cases} \quad (7.8a)$$

$$h_e(x,y) = \begin{cases} h(x,y), & 0 \leq x \leq C-1 \text{ 且 } 0 \leq y \leq D-1 \\ 0, & \text{其他} \end{cases} \quad (7.8b)$$

则输出的降质数字图像为

$$g_e(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m,n) h_e(x-m, y-n) \quad (7.9)$$

式中, $x=0, 1, 2, \dots, M-1$; $y=0, 1, 2, \dots, N-1$ 。

式(7.9)的二维离散退化模型可以用矩阵形式表示,即

$$\mathbf{g} = \mathbf{H}\mathbf{f} \quad (7.10)$$

式中, \mathbf{H} 是 $MN \times MN$ 维矩阵,由 $M \times M$ 个大小为 $N \times N$ 的子矩阵组成,可进一步表示成式(7.11)。将 $g(x,y)$ 和 $f(x,y)$ 中的元素排成列向量, \mathbf{g} 和 \mathbf{f} 成为 $MN \times 1$ 维列向量。

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \cdots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{H}_{M-1} & \cdots & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_1 & \mathbf{H}_0 & \cdots & \mathbf{H}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{M-1} & \mathbf{H}_{M-2} & \mathbf{H}_{M-3} & \cdots & \mathbf{H}_0 \end{bmatrix} \quad (7.11)$$

式中,子矩阵 $\mathbf{H}_j (j=0, 1, 2, \dots, M-1)$ 为分块循环矩阵,大小为 $N \times N$ 。分块矩阵是由延拓函数 $h_e(x,y)$ 的第 j 行构成的,构成方法如下:

$$\mathbf{H}_j = \begin{bmatrix} h_e(j,0) & h_e(j,N-1) & h_e(j,N-2) & \cdots & h_e(j,1) \\ h_e(j,1) & h_e(j,0) & h_e(j,N-1) & \cdots & h_e(j,2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_e(j,N-1) & h_e(j,N-2) & h_e(j,N-3) & \cdots & h_e(j,0) \end{bmatrix} \quad (7.12)$$

将噪声考虑进去,则离散图像退化模型为

$$g_e(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m,n) h_e(x-m, y-n) + n_e(x,y) \quad (7.13)$$

写成矩阵形式为

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} \quad (7.14)$$

上式表明,给定了退化图像 $g(x,y)$ 、退化系统的点扩展函数 $h(x,y)$ 和噪声分布 $n(x,y)$,就可以得到原始图像 f 的估计 \hat{f} 。遗憾的是,上面的线性空间不变模型尽管简单,但实际计算的工作量却十分庞

大。假设图像大小 $M=N$, 则 \mathbf{H} 的大小为 N^4 , 这意味着要解出 $f(x,y)$ 需要解 N^2 个联立方程组。通常有两种解决上述问题的途径:

(1) 通过对角简化分块循环矩阵, 再利用 FFT 快速算法可以大大地降低计算量且能极大地节省存储空间。

(2) 分析退化的具体原因, 找出 \mathbf{H} 的具体简化形式, 如匀速运动造成模糊的 PSF 就可以用简单的形式表示, 这样使复原问题变得简单。

下面将针对式 (7.14) 讨论各种代数复原方法, 它们可能是通过无约束条件而得到原始图像 \mathbf{f} 的估计 $\hat{\mathbf{f}}$, 也可能是约束复原 \mathbf{f} 。

7.2 逆滤波复原

非约束复原指根据对退化系统 \mathbf{H} 和噪声 \mathbf{n} 的了解, 已知退化图像 \mathbf{g} 的情况下, 在一定的最小误差准则下, 得到原始图像 \mathbf{f} 的估计 $\hat{\mathbf{f}}$ 。逆滤波是最早使用的一种无约束复原方法, 成功地应用于航天器传来的退化图像。

由式 (7.14) 可得

$$\mathbf{n} = \mathbf{g} - \mathbf{H}\mathbf{f} \quad (7.15)$$

当对 \mathbf{n} 的统计特性不确定时, 希望对原始图像 \mathbf{f} 的估计 $\hat{\mathbf{f}}$ 应满足这样的条件, 使 $\mathbf{H}\hat{\mathbf{f}}$ 在最小二乘意义上近似于 \mathbf{g} 。也就是说, 希望找到一个 $\hat{\mathbf{f}}$, 使得噪声项的范数

$$\|\mathbf{n}\|^2 = \mathbf{n}^T \mathbf{n} \quad (7.16)$$

最小。即目标函数

$$J(\hat{\mathbf{f}}) = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \quad (7.17)$$

为最小。由极值条件

$$\frac{\partial J(\hat{\mathbf{f}})}{\partial \hat{\mathbf{f}}} = 0 \quad (7.18)$$

得

$$-2\mathbf{H}^T(\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}) = 0 \quad (7.19)$$

在 $M=N$ 的情况下, \mathbf{H} 为方阵且 \mathbf{H} 有逆阵 \mathbf{H}^{-1} , 则

$$\hat{\mathbf{f}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{g} = \mathbf{H}^{-1} \mathbf{g} \quad (7.20)$$

若 \mathbf{H} 已知, 即可由 \mathbf{g} 求出 \mathbf{f} 的最佳估计值 $\hat{\mathbf{f}}$ 。也就是说, 当系统 \mathbf{H} 逆作用于退化图像 \mathbf{g} 时, 可以得到最小平方意义上的非约束估计。

对式 (7.20) 进行傅里叶变换, 则

$$\hat{\mathbf{F}}(u,v) = \frac{\mathbf{G}(u,v)}{\mathbf{H}(u,v)} \quad (7.21)$$

不难看出, 逆滤波法形式简单, 但具体求解的计算量很大, 需要根据循环分块矩阵条件进行简化。该法适用于极高信噪比条件下的图像复原问题, 且降质系统的传递函数 \mathbf{H} 不存在病态性质。将式 (7.7) 带入式 (7.21), 得

$$\hat{\mathbf{F}}(u,v) = \mathbf{F}(u,v) + \frac{\mathbf{N}(u,v)}{\mathbf{H}(u,v)} \quad (7.22)$$

可见当信噪比较低且当 \mathbf{H} 等于 0 或接近于 0 时, 上式后一项将在图像估计中其主导作用。还原的

图像将变得无意义。这时需要人为地对传递函数进行修正,以降低由于传递函数病态而造成的恢复不稳定性。这种处理方法被称为伪逆滤波。第1章中的图1.11是对微装配准备对接模糊图像进行逆滤波方法进行复原的实例,可见这一恢复方法在一定条件下的有效性。

7.3 约束复原

约束复原除了对降质系统的PSF有所了解外,还需要对原始图像和外加噪声的特性有先验知识。根据不同领域的要求,有时需要对 f 和 n 做一些特殊的规定,使处理得到的图像满足某些条件。

7.3.1 约束复原的基本原理

在约束最小二乘法复原问题中,令 Q 为 f 的线性算子,要设法寻找一个最优估计 \hat{f} ,使形式为 $\|Q\hat{f}\|^2$ 的、服从约束条件 $\|g - H\hat{f}\|^2 = \|n\|^2$ 的函数最小化。该最小化问题,可利用拉格朗日乘子法进行处理,即将约束条件表示为 $\alpha(\|g - H\hat{f}\|^2 - \|n\|^2)$,然后加上函数 $\|Q\hat{f}\|^2$,其中 α 为一常数,称为拉格朗日乘子。也就是说,要寻找一个 \hat{f} ,使下面的目标函数(准则函数)为最小:

$$J(\hat{f}) = \|Q\hat{f}\|^2 + \alpha(\|g - H\hat{f}\|^2 - \|n\|^2) \quad (7.23)$$

令 $\frac{\partial J(\hat{f})}{\partial \hat{f}} = 0$,得到 f 的最佳估值 \hat{f} 为

$$\hat{f} = (H^T H + \gamma Q^T Q)^{-1} H^T g \quad (7.24)$$

式中, $\gamma = \alpha^{-1}$ 。式(7.24)是本节讨论的最小二乘法滤波复原的基础,问题的核心是如何选择一个合适的变换矩阵 Q 。 Q 的形式不同,可得到不同类型的最小二乘法滤波复原方法。如选用图像 f 和噪声 n 的自相关矩阵 R_f 和 R_n 表示 Q ,就可得到维纳滤波复原方法。

7.3.2 维纳滤波方法

要掌握图像 f 和噪声 n 的精确先验知识是困难的,一种较为合理的假设是将它们近似地视为平稳随机过程。假设 R_f 和 R_n 为 f 和 n 的自相关矩阵,其定义为

$$R_f = E\{ff^T\} \quad (7.25a)$$

$$R_n = E\{nn^T\} \quad (7.25b)$$

式中, $E\{\cdot\}$ 代表数学期望运算。

定义 $Q^T Q = R_f^{-1} R_n$,代入式(7.24)得

$$\hat{f} = (H^T H + \gamma R_f^{-1} R_n)^{-1} H^T g \quad (7.26)$$

假设 $M = N$, S_f 和 S_n 分别为图像信号和噪声的功率谱,则

$$\begin{aligned} \hat{F}(u, v) &= \left[\frac{H^*(u, v)}{|H(u, v)|^2 + \gamma[S_n(u, v)/S_f(u, v)]} \right] G(u, v) \\ &= \left[\frac{1}{H(u, v)} \cdot \frac{|H(u, v)|^2}{|H(u, v)|^2 + \gamma[S_n(u, v)/S_f(u, v)]} \right] G(u, v) \end{aligned} \quad (7.27)$$

式中, $u, v = 0, 1, 2, \dots, N-1$, $|H(u, v)|^2 = H^*(u, v)H(u, v)$ 。

分几种情况对式 (7.27) 做如下分析。

(1) 如果 $\gamma = 1$, 系统函数 $H_w(u, v)$ 是维纳滤波器的传递函数, 即

$$H_w(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + S_n(u, v) / S_f(u, v)} \quad (7.28)$$

与逆滤波相比, 维纳滤波器对噪声的放大有自动抑制作用。如果无法知道噪声的统计性质, 但可大致确定 $S_n(u, v)$ 和 $S_f(u, v)$ 的比值范围, 式 (7.27) 可以用下式近似

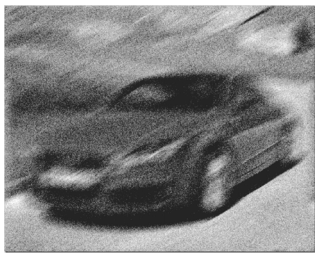
$$\hat{F}(u, v) \approx \left[\frac{H^*(u, v)}{|H(u, v)|^2 + K} \right] G(u, v) \quad (7.29)$$

式中, K 表示噪声对信号的频谱密度之比。

(2) 如果 $\gamma = 0$, 系统变成单纯的去卷积滤波器, 系统的传递函数即为 H^{-1} 。另外一个等效的情况是, 尽管 $\gamma \neq 0$ 但无噪声影响, $S_n(u, v) = 0$, 复原系统亦为理想的逆滤波器, 可以视为维纳滤波器的一种特殊情况。

(3) 若 γ 为可调整的其他参数, 此时为参数化维纳滤波器。一般地, 可以通过选择 γ 的数值来获得所需要的平滑效果。 $H(u, v)$ 由点扩展函数确定, 而当噪声是白噪声时, $S_n(u, v)$ 为常数, 可通过计算一幅噪声图像的功率谱 $S_g(u, v)$ 求解。由于 $S_g(u, v) = |H(u, v)|^2 S_f(u, v) + S_n(u, v)$, 所以 $S_f(u, v)$ 可以求得。研究表明, 在同样的条件下, 单纯去卷积的复原效果最差, 维纳滤波器会产生超过人眼所希望的低通效应; $\gamma < 1$ 的参数化维纳滤波器的图像复原效果较好。

如果满足平稳随机过程的模型和变质系统是线性的两个条件, 那么维纳滤波器将会取得较为满意的复原效果。但是当信噪比很低的情况下, 复原结果还不能令人满意, 那么这可能是由于以下一些因素造成的。



(a) 加性噪声污染和运动模糊图像



(b) 复原后的图像

图 7.2 去卷积维纳滤波复原方法

(1) 维纳滤波器是假设线性系统。但实际上, 图像的记录和评价图像的人类视觉系统往往都是非线性的。

(2) 维纳滤波器是根据最小均方误差准则设计的滤波器, 这个准则不见得与人类视觉判决准则相符合。

(3) 维纳滤波器是基于平稳随机过程的模型, 实际存在的千奇百怪的图像并不

一定都符合这个模型。另外, 维纳滤波器只利用了图像的协方差信息, 可能还有大量的有用信息没有充分利用。

图 7.2 通过实例说明了去卷积维纳滤波复原方法的有效性。

MATLAB 提供了具有维纳滤波的 `deconvwnr` 函数。函数的一般形式是

$$J = \text{deconvwnr}(I, \text{PSF}, \text{NSR}) \quad \text{或} \quad J = \text{deconvwnr}(I, \text{PSF}, \text{NCORR}, \text{ICORR})$$

其中, I 是原始图像, J 是去模糊的图像。NSR 是噪声信号功率比, 默认值为 0, 表示无噪声的情况。NCORR 和 ICORR 表示噪声和原始图像的自相关函数。

下面给出一个在实际中经常会遇到的运动模糊复原问题。在飞机、宇宙飞行器、交通视频监控系统 (车速常常超过 60 km/小时) 等的成像系统与目标有高速相对运动所拍摄的照片, 由于摄取镜头在曝光瞬间 T 的偏移, 会产生匀速直线运动的模糊。设在曝光时间 T 内, 在 x 和 y 方向的运动位移分别是 a 和 b , 可以推导出此时的退化函数为

$$H(u,v) = \frac{T}{\pi(ua+vb)} \sin[\pi(ua+vb)] e^{-j\pi(ua+vb)} \quad (7.30)$$

例 7.1 说明了采用维纳滤波复原的具体实现方法。

【例 7.1】原始图像如图 7.3 (a) 所示, 使用函数 `deconvwnr` 对图 7.3 (b) 所示的无噪声模糊图像进行复原重建, 观察所得结果, 并将不同 PSF 产生的复原效果进行比较。

【解】首先假设真实的 PSF 是由运动形成的, 采用 `PSF = fspecial('motion', LEN, THETA)` 产生一个反映匀速直线运动的二维滤波器。以水平线作为 0 角度基准, 按照逆时针方向, 摄像机按 `THETA` 角方向运动 `LEN` 个像素。默认参数值为 `LEN=9, THETA=0`。

程序代码如下:

```
I = imread('flower.gif');
len=28;
theta=14;
PSF=fspecial('motion',len,theta);
blurred = imfilter(I,PSF,'circular','conv');
%读入无噪声模糊图像并命名为 blurred
len=28;
theta=14;
wnr1=deconvwnr(blurred, PSF); %维纳滤波复原图像
figure, imshow(wnr1);
```



(a) 原始图像



(b) 无噪声模糊图像

图 7.3 原始图像及无噪声模糊图像

复原结果如图 7.4 (a) 所示。实际应用过程中, 真实的 PSF 通常是未知的, 需要根据一定的先验知识对它进行估计, 再将估计值作为参数进行图像复原。图 7.4 分别显示了使用较“长”和较“陡峭”的 PSF 后所产生的复原效果, 由此可见 PSF 的重要性。部分程序代码如下:

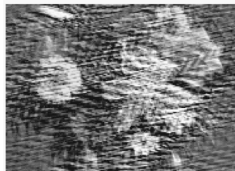
```
wnr2=deconvwnr(blurred, fspecial('motion',2*len,theta)); %长 PSF
wnr3=deconvwnr(blurred, fspecial('motion',len,2*theta)); %陡峭 PSF
```



(a) 使用真实的 PSF 复原



(b) 使用较“长”的 PSF 复原



(c) 使用较“陡峭”的 PSF 复原

图 7.4 不同 PSF 产生的复原效果比较

7.3.3 平滑度约束最小平方滤波

使用逆滤波器一类的方法进行图像复原时, 由于退化算子 H 的病态性质, 导致在零点附近数值起伏过大, 使复原后的图像产生了人为的噪声和边缘 (振铃)。通过选择合理的 Q (高通滤波器), 并对 $\|Qf\|^2$ 进行优化, 可将这种不平滑性降低至最小。使某个函数的二阶导数最小 (如 Q 选用拉普拉斯算子形式表示), 可以推导出以平滑度为基础的约束最小二乘方复原方法。

图像增强的拉普拉斯算子 $\nabla^2 f = \left(\frac{\partial}{\partial x^2} + \frac{\partial}{\partial y^2} \right) f$ 具有突出边缘的作用, 然而 $\iint \nabla^2 f dx dy$ 则恢复了图像的平滑性。因此, 在做图像复原时可将其作为约束。在离散情况下, 拉普拉斯算子 ∇^2 可用下面的 3×3 模板来近似

$$p(x, y) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (7.31)$$

利用 $f(x, y)$ 与上面的模板算子进行卷积可进行高通卷积运算。具体实现时, 可利用添零延拓 $f(x, y)$ 和 $p(x, y)$ 成为 $f_e(x, y)$ 和 $p_e(x, y)$ 来避免交叠误差。在式 (7.23) 中, \mathbf{Q} 对应于高通卷积滤波运算, 在 $\|\mathbf{g} - \mathbf{H}\mathbf{f}\| = \|\mathbf{n}\|^2$ 约束条件下, 最小化 $\|\mathbf{Q}\mathbf{f}\|^2$ 。可以证明, 这时复原 $\hat{\mathbf{f}}$ 的频率域表达式为

$$\hat{F}(u, v) = \left[\frac{H^*(u, v)}{|H(u, v)|^2 + \gamma P(u, v)} \right] G(u, v) \quad (7.32)$$

式中, $u, v = 0, 1, \dots, N-1$, H^* 为 H 的共轭矩阵且 $|H(u, v)|^2 = H^*(u, v)H(u, v)$ 。 γ 的取值控制对所估计图像所加平滑性约束的程度。 $P(u, v)$ 为用 \mathbf{Q} 实现的高通滤波器的传递函数, 决定了不同频率所受平滑性影响的程度。对于拉普拉斯算子有

$$P(u, v) = -4\pi^2(u^2 + v^2) \quad (7.33)$$

MATLAB 提供了在调用维纳滤波的 `deconvwn` 函数、平滑度约束最小二乘滤波的 `deconvreg` 函数前, 降低振铃影响的 `edgetaper` 函数。该函数的输出图像 J 降低了上述算法中由离散傅里叶变换引起的振铃影响。函数的一般形式是

$$J = \text{edgetaper}(I, \text{PSF})$$

`edgetaper` 使用规定的点扩展函数对图像 I 进行模糊操作。

`deconvreg` 函数提供了使用平滑约束最小二乘滤波算法对图像去卷积的功能。调用格式如下

$$[J \text{ LAGRA}] = \text{deconvreg}(I, \text{PSF}, \text{NP}, \text{LRANGE}, \text{REGOP})$$

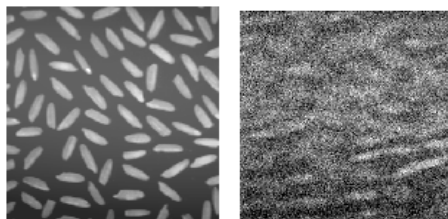
其中, I 假设为真实场景图像在 PSF 的作用下并附加噪声的图像, NP 为噪声强度, J 为去模糊的复原图像。 LRANGE (拉普拉斯算子的搜索范围)、 REGOP (约束算子) 为改善复原效果的可选参数。 LRANGE 指定搜索最佳拉普拉斯算子的范围, 默认值为 $[10^{-9}, 10^9]$ 。返回

值 LAGRA 为在搜索范围的拉格朗日乘子。如果 LRANGE 为标量, 则该算法假定 LAGRA 已经给定且等于 LRANGE , 因而 NP 值可以不予考虑。 REGOP 的默认值为平滑约束拉普拉斯算子。例 7.2 说明了采用平滑约束的最小二乘复原的具体实现方法。

【例 7.2】 原始图像见图 7.5 (a), 对图 7.5 (b) 给出的有噪声模糊图像使用最小二乘滤波方法进行复原重建, 要求尽量提高重建图像的质量。

【解】 程序代码如下:

```
I = imread('rice.tif');
len=28;
theta=14;
PSF=fspecial('motion',len,theta);
blurred = imfilter(I,PSF,'circular','conv'); %生成模糊图像 blurred
v = 0.02;
```



(a) 原始图像 (b) 有噪声模糊图像

图 7.5 原始图像及其有噪声模糊化图像

```

blurrednoisy = imnoise(blurred,'gaussian',0,v);
%读入有噪声模糊图像 blurrednoisy
V=0.02;
NP=V*prod(size(I)); %计算噪声强度
Edged=edgetaper(blurrednoisy,PSF);
%使用 edgetaper 函数对图像进行预处理,以消除图像的不连续边界
[reg1 LAGRA]=deconvreg(Edged,PSF,NP);
figure,subplot(2,2,1), imshow(reg1); %小 NP
reg2=deconvreg(Edged,PSF,NP*1.2);
subplot(2,2,2), imshow(reg2); %大 NP
reg3=deconvreg(Edged,PSF,[],LAGRA);
figure,subplot(2,2,1), imshow(reg3); %小范围搜索
reg4=deconvreg(Edged,PSF,[],50*LAGRA);
subplot(2,2,2), imshow(reg4); %大范围搜索
REGOP=[1 -2 1];
reg5=deconvreg(Edged,PSF,[],LAGRA,REGOP);
figure,subplot(2,2,1),imshow(reg5); %平滑约束复原

```

不同的复原图像效果比较见图 7.6、图 7.7 和图 7.8。通过这些图像,可以分析各参数对图像复原质量的影响。实际应用中,读者可以根据这些经验来选择最佳的参数进行图像复原。



(a) 小 NP



(b) 大 NP

图 7.6 不同信噪比复原结果比较



(a) 小范围搜索



(b) 大范围搜索

图 7.7 不同拉普拉斯算子搜索范围复原效果比较



图 7.8 平滑约束复原效果

7.4 非线性复原方法

前面介绍的复原方法属于经典复原滤波器,其显著特点是约束方程和准则函数中的表达式都可以改为矩阵乘法。这些矩阵都是分块循环矩阵,从而可以实现对角化。下面介绍的方法则属于非线性复原方法,所采用的准则函数都不能进行对角化,因而线性代数的方法在这里是不适用的。

设 S 是非线性函数，当考虑图像的非线性退化时，图像的退化模型可以表示为

$$g(x, y) = S[b(x, y)] + n(x, y) \quad (7.34a)$$

$$b(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, \alpha; y, \beta) f(\alpha, \beta) d\alpha d\beta \quad (7.34b)$$

7.4.1 最大后验复原

与维纳滤波类似，最大后验复原也是一种统计方法。将原图像 $f(x, y)$ 和退化图像 $g(x, y)$ 都看成是随机场，在已知 $g(x, y)$ 的情况下，求出后验概率 $P(f(x, y)|g(x, y))$ 。根据贝叶斯判决理论可知， $P(f|g)P(g) = P(g|f)P(f)$ 。最大后验复原法要求 $\hat{f}(x, y)$ 使式 (7.35) 最大

$$\max_f P(f|g) = \max_f \frac{P(g|f)P(f)}{P(g)} = \max_f P(g|f)P(f) \quad (7.35)$$

最大后验图像复原方法将图像视为非平稳随机场，把图像模型表示成一个平稳随机过程对于一个不平稳的均值做零均值高斯起伏。可以用迭代法求出式 (7.35) 的最佳值。将经过多次迭代、收敛到最后的解作为复原的图像。一种可迭代序列为

$$\hat{f}_{k+1} = \hat{f}_k - h * S_b \{ \sigma_n^{-2} [g - S(h * \hat{f}_k)] - \sigma_f^{-2} (\hat{f}_k - \bar{f}) \} \quad (7.36)$$

式中， k 为迭代次数， $*$ 代表卷积。 S_b 是由 S 的导数构成的函数， σ_f^2 和 σ_n^2 分别为 f 和 n 的方差， \bar{f} 是随空间而变的均值，可视为常数。

式 (7.36) 表明，一个图像的复原可以通过一个序列的卷积来估算，即使 S 是线性的情况下也是适用的，通过人机交互的手段，我们可以在完全收敛前选择一个合适的解。

7.4.2 最大熵复原

1. 正性约束条件

光学图像的数值总为正值，而逆滤波器等线性图像复原可能产生无意义的负输出，这些输出将导致在图像的零背景区域产生一些假的波纹。因此，将复原后的图像 $\hat{f}(x, y)$ 约束为正值是合理的假设。

2. 最大熵复原原理

由于反向滤波器法的病态性，复原出的图像经常具有灰度变换较大的不均匀区域。最小二乘类约束复原方法是最小化的一种反映图像不均匀性的准则函数。最大熵复原方法则是通过最大化某种反映图像平滑性的准则函数来作为约束条件，以解决图像复原中反向滤波法存在的病态问题。

在“图像编码与压缩”一章，我们已经介绍了图像灰度熵的概念。由于概率 $P(k)$ ($k = 0, 1, \dots, M-1$) 介于 $0 \sim 1$ 之间，因此最大熵的范围为 $0 \sim \ln M$ ， H 不可能出现负值，所以最大熵准则能自动地引向全正值的输出结果。

在图像复原中，一种基本的图像熵被定义为

$$H_f = - \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \ln f(m, n) \quad (7.37)$$

最大熵 (Maximum Entropy, ME) 复原的原理是将 $f(x, y)$ 写成随机变量的统计模型，然后在一定的约束条件下，找出用随机变量形式表示的熵表达式，运用求极大值的方法，求得最优估计解 $\hat{f}(x, y)$ 。最大熵复原的含义是对 $\hat{f}(x, y)$ 的最大平滑估计。

3. Friend 和 Burg 复原方法

最大熵复原常用 Friend 和 Burg 两种方法,这两种方法的基本原理相同,但对模型的假设方法不同,得到的最佳估值 \hat{f} 也不相同。两种最大熵复原都是正性约束条件下的图像复原方法,其复原图像的解 $\hat{f}(x, y)$ 是正值,这与光学图像信号要求为正信号相符合。最大化问题都是用拉格朗日系数来完成的。最大熵复原是对原始图像 f 起平滑作用,实质上得到的最优估计 \hat{f} 是最大平滑估计。

(1) Friend 最大熵复原

Friend 法的图像统计模型是将原始图像 $f(x, y)$ 视为由分散在整个图像平面上的离散的数字颗粒组成的。Friend 最大熵复原的基本原理就是求式 (7.37) 为最大来估计原始图像 $\hat{f}(x, y)$ 。数字图像最大熵的复原问题是求一个图像熵和噪声熵加权之和的极大值问题。

Friend 最大熵复原可用迭代方法求解。应用 Newton-Raphson 迭代法求 N^2-1 个拉格朗日系数,一般只需 8~40 次迭代即可求得。

(2) Burg 最大熵复原

最大熵复原是 Burg 于 1967 年在对地震信号的功率谱估计中提出的。假设图像统计模型将 $f(x, y)$ 视为一个变量 $a(x, y)$ 的平方,它保证了 $f(x, y)$ 是正值,即

$$\hat{f}(x, y) = [a(x, y)]^2 \quad (7.38)$$

Burg 定义的熵与式 (7.37) 定义的熵有所不同,其定义为

$$H_f = - \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \ln f(m, n) \quad (7.39)$$

Burg 最大熵复原的基本原理是通过求式 (7.37) 最大来估计 $\hat{f}(x, y)$ 。Burg 最大熵复原可以得到闭合形式解,不需要迭代算法,因而计算时间较短。但此解对噪声比较敏感,如果原始图像中有噪声存在,复原图像可能会被许多小斑点所模糊。

7.4.3 投影复原

投影复原法是用代数方程组来描述线性和非线性退化系统的。退化系统可用下式描述

$$g(x, y) = D[f(x, y)] + n(x, y) \quad (7.40)$$

式中, D 是退化算子,表示对图像进行某种运算。

投影复原的目的是由不完全图像数据求解式 (7.40),找出 $f(x, y)$ 的最佳估计。该法采用迭代法求解与式 (7.40) 对应的方程组。假设退化算子是线性的,并忽略噪声,则式 (7.40) 可写成如下的方程组

$$\begin{cases} a_{11}f_1 + a_{12}f_2 + \dots + a_{1N}f_N = g_1 \\ a_{21}f_1 + a_{22}f_2 + \dots + a_{2N}f_N = g_2 \\ \dots \\ a_{M1}f_1 + a_{M2}f_2 + \dots + a_{MN}f_N = g_M \end{cases} \quad (7.41)$$

式中, f_i 和 g_j ($i=1, 2, \dots, N$; $j=1, 2, \dots, M$) 分别是原始图像 $f(x, y)$ 和退化图像 $g(x, y)$ 的采样, a_{ij} 为常数。投影复原法可以从几何学角度进行解释。 $f=[f_1, f_2, \dots, f_N]$ 可视为在 N 维空间中的一个向量,而式 (7.41) 中的每一个方程代表一个超平面。下面采用投影迭代法找到 f_i 的最佳估值。

迭代法首先假设一个初始估值 $f^{(0)}(x, y)$, 然后进行迭代运算,第 k 次迭代值 $f^{(k)}(x, y)$ 由其前次迭代值 $f^{(k-1)}(x, y)$ 和超平面的参数决定。可以根据退化图像取初始估值。下一个推测值 $f^{(1)}$ 取 $f^{(0)}$ 在第一个超平面 $a_{11}f_1 + a_{12}f_2 + \dots + a_{1N}f_N = g_1$ 上的投影,即

$$\mathbf{f}^{(1)} = \mathbf{f}^{(0)} - \frac{(\mathbf{f}^{(0)} \cdot \mathbf{a}_1 - g_1)}{\mathbf{a}_1 \cdot \mathbf{a}_1} \mathbf{a}_1 \quad (7.42)$$

式中, $\mathbf{a}_1 = [a_{11}, a_{12}, \dots, a_{1N}]$, 圆点代表向量的点积。

再取 $\mathbf{f}^{(1)}$ 在第二超平面 $a_{21}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + \dots + a_{2N}\mathbf{f}_N = g_2$ 上的投影, 并称之为 $\mathbf{f}^{(2)}$, 依次向下, 直到得到 $\mathbf{f}^{(M)}$ 满足式 (7.39) 中最后一个方程式。这样就实现了迭代的第一个循环。

然后从式 (7.39) 的第一个方程式中开始第二次迭代。即取 $\mathbf{f}^{(M)}$ 在第一个超平面 $a_{11}\mathbf{f}_1 + a_{12}\mathbf{f}_2 + \dots + a_{1N}\mathbf{f}_N = g_1$ 上的投影, 并称之为 $\mathbf{f}^{(M+1)}$, 再取 $\mathbf{f}^{(M+1)}$ 在 $a_{21}\mathbf{f}_1 + a_{22}\mathbf{f}_2 + \dots + a_{2N}\mathbf{f}_N = g_2$ 上的投影……直到式 (7.39) 中的最后一个方程式。这样, 就实现了迭代的第二个循环。按照上述方法不断地迭代, 便可得到一系列向量 $\mathbf{f}^{(0)}, \mathbf{f}^{(M)}, \mathbf{f}^{(2M)}, \dots$ 。可以证明, 对于任意给定的 N, M 和 a_{ij} , 向量 $\mathbf{f}^{(kM)}$ 将收敛于 \mathbf{f} , 即

$$\lim_{k \rightarrow \infty} \mathbf{f}^{(kM)} = \mathbf{f} \quad (7.43)$$

投影迭代方法要求有一个好的初始估值 $\mathbf{f}^{(0)}$ 开始迭代。在应用此法进行图像复原时, 还可以很方便地引进一些先验信息附加的约束条件, 如 $f_i \geq 0$ 或 f_i 限制在某一范围之内, 可改善图像复原效果。

采用迭代算法的图像非线性复原算法还有蒙特卡罗复原法等。蒙特卡罗复原方法的主要思想是: 把图像分成许多细胞, 相当于像素, 同时认为图像的灰度是由颗粒组成的, 每个颗粒具有一定的能量 d_0 , 并假定图像的颗粒总数是已知的, 即图像的总能量是已知的。某一颗粒随机地分配到某一细胞中, 满足一定的判决原则。如果所有的颗粒都分配完了, 那么最后的目标图像也就复原出来了。蒙特卡罗复原方法具有使用灵活、容易满足和实现正值约束及限带等许多约束条件且计算速度快等优点。在无噪声或噪声较小, 点扩展函数和图像数据存在零点以及二值图像的情况下, 可取得较好的复原效果。

7.4.4 同态滤波复原

自然景物的图像是由照明函数和反射函数两个分量的乘积组成的。同态滤波法 (homomorphic filtering) 复原方法是基于图像的乘性结构理论而提出来的。当降质图像是由两个分量相乘得到时, 可先对降质图像取对数, 得到两个相加的分量, 再进行滤波处理, 最后通过指数变换得到复原图像 $\hat{f}(x, y)$ 。

设退化图像 $g(x, y)$ 可以分为两部分的乘积, 即

$$g(x, y) = i(x, y)r(x, y) \quad (7.44)$$

取对数得

$$\log g(x, y) = \log i(x, y) + \log r(x, y) \quad (7.45)$$

设同态滤波器的冲激响应为 $l(x, y)$, 其复原结果为 $\hat{f}(x, y)$ 。同态滤波复原过程如图 7.9 所示。

可见, 同态复原结果 $\hat{f}(x, y)$ 为

$$\hat{f}(x, y) = 10^{l(x, y) * [\log i(x, y) + \log r(x, y)]} \quad (7.46)$$

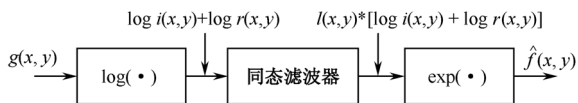


图 7.9 同态滤波器复原

在不考虑相位的情况下, 同态滤波也可用频率域复原方法进行。其复原的准则是估计图像 $\hat{f}(x, y)$ 的功率谱 $S_{\hat{f}}(u, v)$ 与原图像 $f(x, y)$ 的功率谱 $S_f(u, v)$ 相等, 由式 (7.14) 指出的图像退化模型 $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$, 可以推导出同态滤波器的传递函数为

$$|L(u, v)| = \left[\frac{1}{|H(u, v)|^2 + S_n(u, v) / S_f(u, v)} \right] \quad (7.47)$$

同态滤波器的传递函数与维纳滤波器的形式基本相似。如噪声项为零, 其滤波器的传递函数为 $1/H(u, v)$, 这就是逆滤波器。

不难看出,同态滤波技术也可以用于图像增强。采用同态滤波可以实现同态增晰,既能使图像的灰度动态范围压缩,又能使感兴趣的物体图像灰度级扩展。

7.5 盲图像复原

多数复原技术以图像退化的某种先验知识为基础。但是,很多情况下难以确定退化的点扩展函数和噪声的统计特性。盲图像复原法是在没有图像退化必要的先验知识的情况下,对观察的图像以某种方式提取出退化信息,采用盲去卷积算法对图像进行复原。

对具有加性噪声的模糊图像做盲图像复原的方法一般有两种:直接测量法和间接估计法。

7.5.1 直接测量法

用此法复原图像时,通常需要测量图像的模糊脉冲响应和噪声功率谱或协方差函数。在所观察的景物中,往往点光源能直接指示出冲激响应。另外,图像边缘是否陡峭也能用来推测模糊冲激响应。在背景亮度相对恒定的区域内测量图像的协方差,可以估计出观测图像的噪声协方差函数。

7.5.2 间接估计法

间接估计法复原类似于多图像平均法处理。例如,在电视系统中,观测到的第 i 帧图像为

$$g_i(x, y) = f(x, y) + n_i(x, y) \quad (7.48)$$

式中, $f(x, y)$ 是原始图像, $g_i(x, y)$ 是含有噪声的图像, $n_i(x, y)$ 是加性噪声, $i = 1, 2, \dots, N$ 。如果原始图像在 N 帧观测图像内保持恒定,对 N 帧观测图像求和,得到下式

$$\hat{f}(x, y) = \frac{1}{N} \sum_{i=1}^N g_i(x, y) - \frac{1}{N} \sum_{i=1}^N n_i(x, y) \quad (7.49)$$

当 N 很大时,上式右边的噪声项的值趋向于其数学期望 $E\{n(x, y)\}$ 。一般情况下,高斯白噪声在所有 (x, y) 上的数学期望等于零。因此,合理的估计量是

$$f(x, y) = \frac{1}{N} \sum_{i=1}^N g_i(x, y) \quad (7.50)$$

盲图像复原的间接估计法也可以利用时间上平均的概念去掉图像中的模糊。如果有一成像系统,其中相继帧含有相对平稳的目标退化,这种退化是由于每帧有不同的线性位移不变冲激响应 $h_i(x, y)$ 引起的。例如,大气湍流对远距离物体摄影就会产生这种图像退化。只要物体在帧间没有很大移动并且每帧取短时间曝光,那么第 i 帧的退化图像 $g_i(x, y)$ 可表示为

$$g_i(x, y) = f_i(x, y) * h_i(x, y) \quad (7.51)$$

式中, $f_i(x, y)$ 是原始图像, $g_i(x, y)$ 是退化图像, $h_i(x, y)$ 是点扩展函数,*表示卷积, $i = 1, 2, \dots, N$ 。退化图像的傅里叶变换为

$$G_i(u, v) = F_i(u, v) H_i(u, v) \quad (7.52)$$

利用同态处理方法把原始图像的频谱和退化传递函数分开,则可得到

$$\ln [G_i(u, v)] = \ln [F_i(u, v)] + \ln [H_i(u, v)] \quad (7.53)$$

如果帧间退化冲激响应是不相关的,则可得到下面的表达式

$$\sum_{i=1}^N \ln [G_i(u, v)] = N \ln [F_i(u, v)] + \sum_{i=1}^N \ln [H_i(u, v)] \quad (7.54)$$

当 N 很大时, 传递函数的对数和式接近于一恒定值, 即

$$K_H(u, v) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \ln[H_i(u, v)] \quad (7.55)$$

因此, 图像的估计量为

$$\hat{F}_i(u, v) = \exp\left\{-\frac{K_H(u, v)}{M}\right\} \prod_{i=1}^M [G_i(u, v)]^{\frac{1}{M}} \quad (7.56)$$

式中, M 为观测图像的帧数。对上式取傅里叶逆变换就可得到空域估计 $\hat{f}(x, y)$ 。

以上分析没有考虑加性噪声的影响, 否则便无法进行原始图像与点扩散函数的分离处理, 后面的推导也就不成立了。为了解决这一问题, 可以对观测到的每帧图像先进行滤波处理, 消除或降低噪声的影响, 再进行上述处理。

MATLAB 提供了 deconvblind 函数进行盲图像复原。该函数的一般形式是

$$[J, \text{PSF}] = \text{deconvblind}(I, \text{INITPSF}, \text{NUMIT}, \text{DAMPAR}, \text{WEIGHT}, \text{READOUT})$$

该函数采用最大似然算法 (maximum likelihood algorithm) 对模糊图像 I 进行去卷积处理, 返回去模糊的图像 J 和相应的点扩散函数 PSF 。 PSF 是与 INITPSF 大小一样的正值矩阵, 归一化和等于 1。 PSF 复原效果强烈地受到初始值 INITPSF 大小的影响, 而较少受到其元素值的影响, 因而用元素全为 1 的矩阵是一种安全的初值。

为了改进复原效果, 可以选择附加的参数 $\text{NUMIT}, \text{DAMPAR}, \text{WEIGHT}, \text{READOUT}$ 。 NUMIT 是迭代次数, 默认值为 10。 DAMPAR 是指定由图像 I 产生的图像的阈值偏移 (即泊松噪声的标准差, 在该值以下将会发生阻尼现象) 的一个矩阵, 默认值为 0, 表示无阻尼。

WEIGHT 是一个与原始图像 I 大小相同的权矩阵, 反映每个像素在摄取过程中的质量, 如果赋以 0 加权值, 则用来屏蔽差的像素, 而好的像素则被赋以加权值 1。默认值是加权值均为 1 的权矩阵。复原的图像在灰度变化较大的边界或部分存在一定的“环”, 合理地使用 WEIGHT 参数可消除这些环以提高图像的复原质量。

READOUT 是摄取设备的读出噪声方差矩阵。默认值为 0 矩阵, 表示无噪声的情况。

甚至用户提供的对 PSF 约束的附加函数句柄 (FUN) 也可以加入到上述函数中。函数的形式扩展为

$$[J, \text{PSF}] = \text{deconvblind}(\dots, \text{FUN}, \text{P1}, \text{P2}, \dots, \text{PN})$$

其中, $\text{P1}, \text{P2}, \dots, \text{PN}$ 为 FUN 所接受的参数。

【例 7.3】 原始图像见图 7.10 (a), 对图 7.10 (b) 所示的图像进行复原。

【解】 调用 deconvblind 函数进行图像复原时, INITPSF 是一个非常重要的指标。先用较小的 PSF 对图像进行复原, 然后再用一个较大的 PSF 对图像进行复原, 最后与采用真实的 PSF 进行比较。程序代码如下:

```
I = imread('autumn.tif');
len=28;
theta=14;
PSF=fspecial('motion',len,theta);
blurred = imfilter(I,PSF,'circular','conv'); %生成模糊图像 blurred
v = 0.02;
blurrednoisy = imnoise(blurred,'gaussian',0,v);
%读入有噪声的模糊图像并命名为 Noisy
len=9;
theta=12;
PSF=fspecial('gaussian',len,theta); %真实的 PSF
```

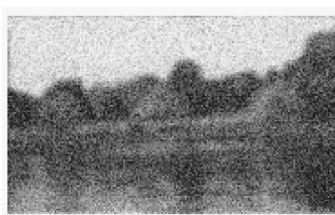
```

UNDERPSF=ones(size(PSF)-4);
%该 PSF 数组的每一维都比真实的 PSF 少 4 个像素
[J1 P1]=deconvblind(Noisy,UNDERPSF);
figure,subplot(2,2,1); imshow(J1);
OVERPSF=padarray(UNDERPSF,[4 4],'replicate','both');
%该 PSF 数组的每一维都比真实的 PSF 多 4 个像素
[J2 P2]=deconvblind(Noisy,OVERPSF);
subplot(2,2,2); imshow(J2);
INITPSF=padarray(UNDERPSF,[2 2],'replicate','both');
[J3 P3]=deconvblind(Noisy,INITPSF);
subplot(2,2,3); imshow(J3);

```



(a) 原始图像



(b) 有噪声模糊图像

图 7.10 原始图像及其有噪声模糊化图像

复原后的图像如图 7.11 所示。



(a) 较小的 PSF



(b) 较大的 PSF



(c) 真实的 PSF

图 7.11 采用不同大小的 PSF 进行图像复原的效果比较

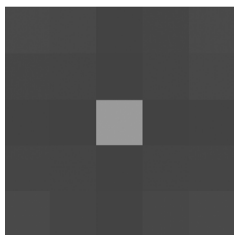
由此可见, PSF 对图像复原质量有非常重要的影响。实际应用中, 可以通过对 PSF 的分析来选择一个合适的 PSF 进行图像复原。例如, 使用以下代码显示图 7.10 返回的重建 PSF 以及真实的 PSF:

```

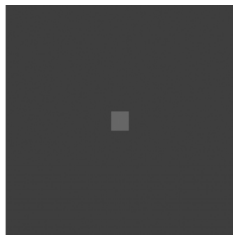
figure,subplot(1,3,1);imshow(P1);
subplot(1,3,2);imshow(P2);
subplot(1,3,3);imshow(P3);

```

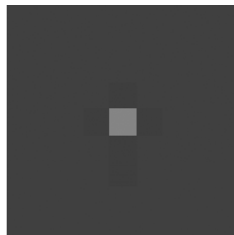
结果如图 7.12 所示。



(a) 较大的 PSF



(b) 较小的 PSF



(c) 真实的 PSF

图 7.12 不同大小 PSF 的图像比较

7.6 几何失真校正

图像在获取过程中, 由于成像系统的非线性、飞行器的姿态变化等原因, 成像后的图像与原景物图像相比, 会产生比例失调, 甚至扭曲。这类图像退化现象称之为几何失真(畸变)。有几何畸变的图像, 不但视觉效果不好, 而且在对图像进行定量分析时提取的形状、距离、面积等数据也不准确。但有时为了表现特殊的图像获取效果, 明显的镜头失真在所难免。如鱼镜头具有焦距短、视场角大的优点, 但用鱼眼摄像机拍摄的图像具有严重的变形失真, 重现时常常需要将其转换成人眼视觉所习惯的透视投影图像。

7.6.1 典型的几何失真

以下以遥感图像为例说明典型的几何失真。由于遥感图像的获取存在着许多不稳定因素, 遥感图像最容易产生几何失真, 一般可分为两类。

1. 系统失真

光学系统、电子扫描系统失真而引起的斜视畸变、枕形、桶形畸变等, 都可能使图像产生几何特性失真。典型的系统失真如图 7.13 所示。

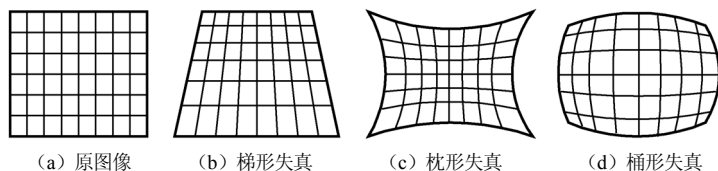


图 7.13 典型的系统几何失真

2. 非系统失真

从飞行器上所获得的地面图像, 由于飞行器的姿态、高度和速度变化引起的不稳定与不可预测的几何失真, 这类畸变一般要根据航天器的跟踪资料 and 在地面设置控制点的办法来进行校正。典型的非系统失真如图 7.14 所示。

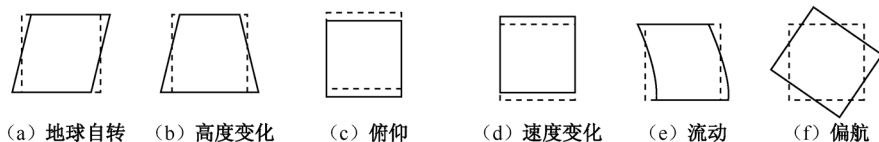


图 7.14 典型的非系统几何畸变

一般来说, 几何畸变校正要对失真的图像进行精确的几何校正, 通常先要确定一幅图像为基准, 然后去校正另一幅图像的几何形状。因此, 几何畸变校正一般分两步来做: 第一步是图像空间坐标的变换; 第二步是重新确定在校正空间各像素点的取值。

7.6.2 空间几何坐标变换

如图 7.15 所示, 空间几何坐标变换指按照一幅标准图像 $f(x, y)$ 或一组基准点去校正另一幅几何失真图像 $g(x', y')$ 。根据两幅图像的一些已知对应点对(控制点对)建立起函数关系式, 将失真图像的 $x'-y'$ 坐标系变换到标准图像 $x-y$ 坐标系, 从而实现失真图像按标准图像的几何位置校正, 使 $f(x, y)$ 中的每一像点都可在 $g(x', y')$ 中找到对应像点。

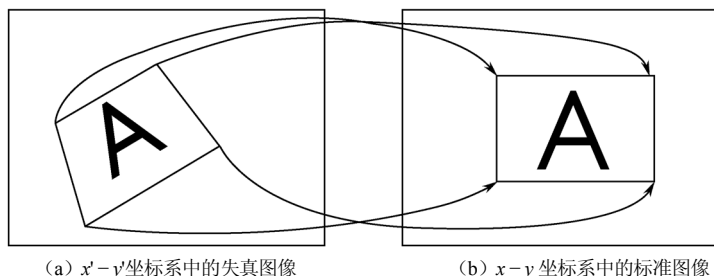


图 7.15 几何位置校正

设原图像用 (x, y) 坐标系, 畸变图像用 (x', y') 坐标系。两个坐标系之间的关系为

$$\begin{cases} x' = h_1(x, y) \\ y' = h_2(x, y) \end{cases} \quad (7.57)$$

几何校正方法可以分为两类: 一类是在 h_1 、 h_2 已知情况下的校正方法, 另一类是在 h_1 、 h_2 未知情况下的校正方法。前者一般通过人工设置标志来进行, 如卫星照片通过人工设置小型平面反射镜作为标志。后者通过控制点之间的空间对应关系建立线性 (如三角形线性法) 或高次 (如二元二次多项式法) 方程组求解式 (7.57) 中坐标之间的对应关系。下面以三角形线性法为例讨论空间几何坐标变换问题。

某些图像, 如卫星所摄天体照片, 对大面积来讲, 图像的几何失真虽然是非线性的, 但在一个小区域内可近似认为是线性的。这时就可将畸变系统和校正系统坐标用线性方程来联系。将标准图像和被校正图像之间的对应点对划分成一系列小三角形区域, 三角形顶点为 3 个控制点, 在三角形区内满足以下线性关系:

$$\begin{cases} x' = ax + by + c \\ y' = dx + ey + f \end{cases} \quad (7.58)$$

解方程组 (7.58), 可求出 a 、 b 、 c 、 d 、 e 、 f 六个系数。用式 (7.57) 可实现该三角形区内其他像点的坐标变换。对于不同的三角形控制区域, 这 6 个系数的值是不同的。

三角形线性法简单, 能满足一定的精度要求, 这是因为它是局部范围内的线性失真去处理大范围内的非线性失真, 所以选择的控制点对越多, 分布越均匀, 三角形区域的面积越小, 则变换的精度越高。但是控制点过多又会导致计算量的增加。

7.6.3 校正空间像素点灰度值的确定

图像经几何位置校正后, 在校正空间中各像素点的灰度值等于被校正图像对应点的灰度值。一般校正后的图像的某些像素点可能分布不均匀, 不会恰好落在坐标点上, 一般映射到输入图像中的非整数位置, 即位于 4 个输入像素之间。由于图像像素的灰度值是离散的, 校正需要对原来在整数点坐标上的像素值进行插值生成连续灰度值, 因此常采用在水平和垂直方向同时进行插值的二维内插法来求得这些像素点的灰度值。另外, 数字图像的缩放是一种常见的操作, 数字图像进行缩放后, 也需要重新计算缩放后图像新的像素坐标和相应的灰度, 然后在插值空间域重新采样以获得缩放后图像像素的灰度值。

经常使用的方法有最近邻点法 (nearest neighbor interpolation)、双线性插值法 (bilinear interpolation)、三次卷积法 (又叫立方卷积法, cubic convolution), 其中三次卷积法精度最高, 但计算量也较大。下面介绍前两种方法。实际工作中, 往往综合需要的插值的光滑程度等效果、插值计算量、占用计算机内存等因素选择合适的插值方法。效果越好, 往往需要计算机硬件的开销也越大。它是在二维三次卷积插值法的基础上, 利用三次样条插值法的插值函数具有三阶连续导数的性质而得来的。

1. 最近邻点法

该法中输出像素的灰度值取与该像素点相邻的4个映射点中距离最近的邻点输入灰度值作为该点的灰度值,属于零阶插值法。如图7.16所示。显然,最近邻点法计算十分简单,速度最快,在许多情况下其结果也可令人接受。但精度不高,同时校正后的图像亮度不够平滑,即当图像中包含像素之间灰度级有变化的细微结构时,会在图像中产生方块或锯齿效应,有明显的 discontinuity。

2. 双线性插值法

又称为一阶插值算法,是对最近邻点法的一种改进。如图7.17所示,设标准图像像素坐标 (x_0, y_0) 对应于失真图像像素坐标 (x'_0, y'_0) ,而 (x'_0, y'_0) 点周围4个点的坐标分别为 (x'_1, y'_1) 、 (x'_1+1, y'_1) 、 (x'_1, y'_1+1) 和 (x'_1+1, y'_1+1) ,用 (x'_0, y'_0) 点周围4个邻点的灰度值加权内插作为灰度校正值 $f(x_0, y_0)$,则

$$f(x_0, y_0) = g(x'_0, y'_0) = (1-\alpha)(1-\beta)g(x'_1, y'_1) + \alpha(1-\beta)g(x'_1+1, y'_1) + (1-\alpha)\beta g(x'_1, y'_1+1) + \alpha\beta g(x'_1+1, y'_1+1) \quad (7.59)$$

式中, $\alpha = |x'_0 - x'_1|$, $\beta = |y'_0 - y'_1|$ 。

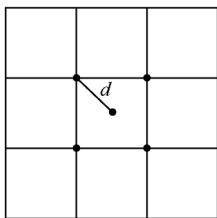


图 7.16 最近邻点法

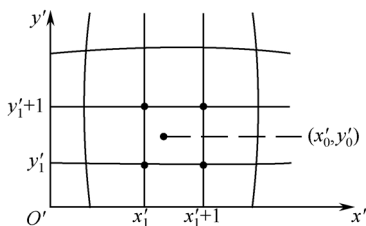


图 7.17 双线性插值法几何校正

与最近邻点法相比,双线性插值法几何校正灰度连续,结果一般满足要求,但占用内存较多,计算量较大且具有低通特性,在灰度突变处的斜率会发生改变,导致图像轮廓模糊。

如果要进一步改善图像质量,可以选用三次卷积法。该算法考虑一个浮点坐标周围的16个邻点,通过分段 Hermite 插值方法计算灰度值,其插值数据和导数都是连续的。当然该方法付出的代价是需要更多的内存和运算时间。

还可以选择三次样条插值法、基于小波的插值法、根据影射点邻域图像复杂度自适应地调整插值计算权值的插值算法等。

MATLAB 图像处理工具箱提供了图像缩放的函数。

$$J = \text{imresize}(I, \text{SCALE}, \text{METHOD})$$

其中, I、J 分别为原图像和缩放后的图像; SCALE 为水平和垂直缩放尺度,小于1为缩小,大于1为放大; METHOD 为插值法,可选'nearest' (默认值)、'bilinear'、'bicubic'。

【例 7.4】原始图像见图 7.18 (a),对该图像分别以 0.5 (1/2) 和 1.414 ($\sqrt{2}$) 进行缩放。

【解】调用 imresize 函数对尺寸为 198×135 原始图像 onion.png 采用不同的方法进行缩放。程序代码如下:

```
%以 0.5 的尺度对原始图像进行缩小
I = imread('onion.png');
I1 = rgb2gray(I);
J1_ZoomOut = imresize(I1,0.5); %默认为最邻近点法压缩
figure(1);
subplot(1,2,1); imshow(I1);
```



```
subplot(1,2,2); imshow(J1_ZoomOut);

%以 1.414 的尺度对原始图像进行放大
J1_ZoomIn = imresize(I1,1.414,'bilinear'); %双线性插值法放大
J2_ZoomIn = imresize(I1,1.414,'bicubic'); %双立方插值法放大
figure(2);
subplot(1,2,1); imshow(J1_ZoomIn);
subplot(1,2,2); imshow(J2_ZoomIn);
```

结果如图 7.18 所示。由图可见，图像尺寸的压缩（仅为原图像像素数的 $1/4$ ）导致图像空间率下降，平滑的灰度出现马赛克现象，边缘有锯齿状。图像放大后的图像由于插值的点较多（是原图像的 1.414^2 倍，即 2 倍），因此在一定的空间分辨率下，可以显示放大 2 倍的图像而不出现明显的灰度失真。

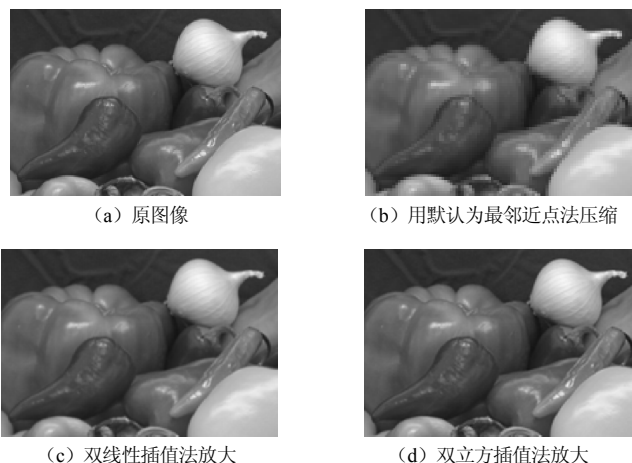


图 7.18 不同插值法对图像的缩放效果

在医学、遥感、安防、视频服务等场合，人们经常期望得到高分辨率图像。但高分辨率图像的获取往往受到传感器、光学制造技术和用户对设备性价比要求的限制。图像超分辨率（Super-Resolution, SR）技术是指通过信号处理硬件或软件的方法由一幅或一系列低分辨率（Low Resolution, LR）图像恢复出高分辨率（High Resolution, HR）图像。图像插值与 SR 是相关的技术，其目的都是要增加单幅图像的尺寸。但由于对单幅图像插值不能恢复在 LR 采样过程中损失的高频部分，故从一幅近似的 LR 图像来提高图像的质量是十分有限的，因此图像插值方法不等同于 SR 技术。有效的途径是基于同一场景的不同的观察数据的融合实现 SR 复原。

普通的数码相机由于光学镜头的生产工艺等原因，拍摄的图像会出现较为严重的非线性几何畸变。如图 7.19 所示，通过校正，可以较好地复原原始场景图像。

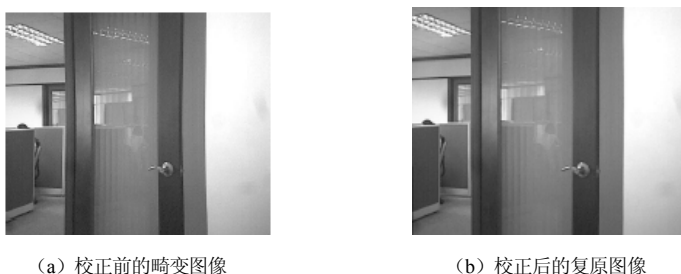


图 7.19 非线性几何畸变校正实例

7.6.4 鱼眼图像校正方法简介

鱼眼镜头又称全景镜头，其前镜片凸出，酷似鱼的眼睛。国家标准 GB/T 13964-92 中定义“鱼眼镜头是视场角超过 140° 的照相镜头，有的几乎达到 180° ，有的甚至超过 220° ”。通过鱼眼镜头拍摄得到的照片是一组发生明显桶型畸变的图片。以上两种对鱼眼镜头的解释有着明显的差异，综合而言，所谓“鱼眼”，即视场角极大，镜头像鱼眼那样外凸。用此类镜头拍摄的照片典型特征是具有明显的桶形畸变。至于视场角究竟是多少度才算鱼眼镜头，以国家标准“超过 140° ”为准。众所周知，焦距越短，视角越大，因光学原理产生的变形也就越强烈。为了达到 180° 的超大视角，鱼眼镜头的设计者不得不作出牺牲，即允许这种变形（桶形畸变）的合理存在。其结果是除了画面中心的景物保持不变，其他本应水平或垂直的景物都发生了相应的变化。也正是这种强烈的视觉效果为那些富于想像力和勇于挑战的摄影者提供了展示个人创造力的机会。

由于鱼眼镜头具有焦距短、视场角大的优点，已被广泛应用在视频监控、智能交通系统、医学成像、数码特技等全方位视觉系统领域。但是鱼眼镜头拍摄的图像严重变形，如果想利用这些变形的信息，就需要将这些图像校正为符合人们视觉习惯的透视投影图像。鱼眼镜头标定算法是一类精确恢复的方法，在建立鱼眼镜头变形模型的基础上，考虑到鱼眼镜头成像的各种畸变类型，如常见的径向变形、离心变形、薄棱镜变形等，建立精确的鱼眼镜头成像模型。对整幅鱼眼图像的校正，一般有两种方法。

第一种方法是从鱼眼成像的两种投影模型——球面投影模型和抛物面投影模型来分析。但用该模型计算过于复杂，实际很少应用。

第二种方法是从 2D 和 3D 空间进行鱼眼图像变形校正，该类方法包括经度坐标校正、多项式坐标变换，以及极半径映射等。经度坐标校正算法较简单，但校正结果存在一定的失真。投影转换算法是将鱼眼图像转换成透视投影图像，也可认为是鱼眼图像半径和入射角间关系的映射。

鱼眼图像包含的信息很丰富，然而实际检测中人们可能并不是特别关心整个全景区域，而是对某个区域感兴趣，这就需要针对感兴趣区域进行畸变校正，使观察重点位于感兴趣区域中心。因此，鱼眼图像部分感兴趣区域校正的研究显得尤为重要。图 7.20（a）是一张俯视拍摄鱼眼畸变图像，图像中的主要内容集中在圆内。人眼对这种图像难以接受，所以必须对它进行校正。校正主要是对图像中的圆形鱼眼中的像素进行坐标变换，使图像转变为一个平视的矩形图像。手动选择感兴趣校正区域，可以很快计算出圆心横纵坐标及半径。有了中心点坐标及内外圆半径，就可以根据坐标对应关系将鱼眼图像映射到矩形区域上。从图 7-20（b）可见，原图 7.20（a）中的失真已经消失，图像变成人眼可以轻易看清的图像。通过与原失真图像对比可以看出，失真图像中除中心点及其周围一小片区域外，其他景物均可以在校正后的图像中找到，并且更加直观。



（a）俯视拍摄鱼眼畸变图像



（b）校正后的俯视拍摄图像

图 7.20 俯视拍摄鱼眼畸变图像校正实例

7.7 图像修复技术简介

由于各种自然或人为的原因,数码相片、影像资料等多媒体载体往往会局部产生破损,这使得资料的完整性遭到了破坏,不符合人类视觉的需要。数字图像修复技术旨在利用这些破损资料中残留的信息,通过某种算法或操作方法,按照一定的修复顺序,把已知信息传播到破损区域,以完成填充破损区域的过程。修复的结果不仅要强调图像作为一个整体所具有的完整美观,同时也要兼顾对于人类视觉的合理性和连贯性。

目前,数字图像修复技术已被广泛推广,特别在以下的几个方面获得了较大进展:① 修复旧损艺术品、照片、电影胶片,如图 7.21 所示;② 移除图像中的指定目标,如图 7.22 所示;③ 移除图像中的字幕、划痕,如图 7.23 所示;④ 图像压缩及无线传输中丢失帧的修复;⑤ 超分辨率方面的研究,如图 7.24 所示。



(a) 破损的古艺术品



(b) 修复过的古艺术品

图 7.21 古油画修复图



(a) 原完好照片



(b) 指定人物移除后结果

图 7.22 资料照片指定目标移除效果图



(a) 包含字幕的图像

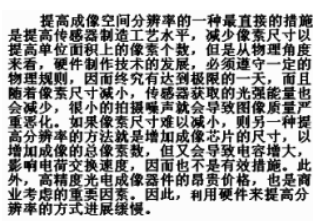


(b) 去除字幕后的图像

图 7.23 图像修复字幕去除效果



(a) 一幅降质的文档图像



(b) 经过图像修复后的超分辨率结果

图 7.24 图像修复在超分辨率研究方面的应用

为了实现图像自动修复的功能并提高修复速度和质量,出现了不少优秀的模型及方法,主要有:

① 基于变分和偏微分方程的图像修复方法。M. Bertalmio 等人于 2000 年的 SIGGRAPH 会议上,首先提出了数字图像修复 (Digital image inpainting) 的概念,并引入了基于偏微分方程 (Partial Differential Equation, PDE) 的算法,即 BSCB 模型。② 基于等照度线扩散的图像修复方法。由于基于偏微分方程的修复算法十分耗时, Oliveira 等人提出了一种快速修复算法,通过高斯卷积核,对破损区域进行重复卷积,将破损区域的边缘信息扩散到破损区域的内部。③ 基于纹理合成的图像修复方法。Efros 和 Leung 提出了根据单个像素点合成、进行非参数采样的纹理合成算法,此算法选用了马尔可夫随机场 (Markov Random Field, MRF) 模型,在处理结构性纹理图像的合成效果比较好。④ 混合修复方法。对于结构信息和纹理信息都比较复杂的图像,无论单使用基于偏微分方程的方法还是单使用基于纹理合成的方法,都无法获得较为满意的修复结果。因此,出现了各取二者所长、两种方法混合修复的新算法。⑤ 基于变换域方法。基于傅里叶变换和小波变换等方法,也通过利用图像的稀疏表示 (Sparse Representation) 并结合曲波变换 (Curvelet Transformation) 和 DCT 等来实现对破损图像进行修复。

对于图像修复结果的评价,一般可以从两方面进行考量:一是修复的效率,通常直接用运算时间来衡量;二是修复的效果,修复的效果由于图像修复本身具有病态特性,使得简单用数学模型来具体描述比较困难,同时也难以找到统一的标准去衡量效果的优劣。再加上图像本身以及破损区域复杂度的影响,不同类型的图像修复算法往往采用不同的衡量标准。图 7.27 所示为缺损修复测试的实例。

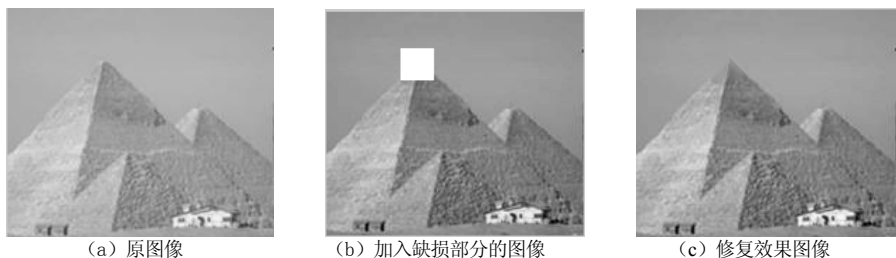


图 7.25 缺损修复测试

7.8 实验：图像复原

1. 实验目的

- (1) 了解图像复原的原理。
- (2) 掌握常用图像复原方法。

2. 实验主要仪器设备

- (1) 台式机或笔记本电脑。
- (2) MATLAB 软件 (含图像处理工具箱)。

3. 实验原理

- (1) 了解不同条件下的图像退化成因和处理方法。
- (2) 根据退化的数学模型对退化图像进行处理,恢复原始图像。

4. 实验内容

- (1) 利用维纳滤波对有噪声模糊图像进行复原。
- (2) 比较直接维纳滤波、设置信噪比参数、设置噪声和自相关函数这三种情况下的图像复原效果。

5. 实验步骤

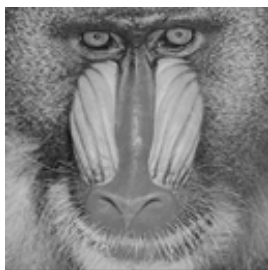
- (1) 选择图 7.26 所示的原始图像作为测试图像。
- (2) 设计图像复原算法。



(a) eight.tif



(b) hill.tif



(c) mandrill.tif

图 7.26 供图像复原的测试图像

- (3) 设置维纳滤波中的相关参数。
- (4) 显示原始图像和复原后的重建图像。
- (5) 通过修改这些参数来观察参数对图像复原效果的影响。

6. 实验报告要求

- (1) 说明图像复原原理与算法。
- (2) 原始图像和复原后的重建图像。
- (3) 对于同一幅图像，比较直接维纳滤波、设置信噪比参数、设置噪声和自相关函数这三种情况下的图像复原效果。
- (4) 回答思考题。
- (5) 心得和体会。

7. 预习要求

- (1) 了解图像复原的基本原理和算法。
- (2) 掌握 IPT 中有关图像复原的函数。
- (3) 画出维纳滤波复原方法的流程框图，编写相应的程序。

8. 思考题

- (1) 比较维纳滤波复原方法与其他复原方法的适用范围。
- (2) 为了尽量提高图像重建质量，应该如何设置其中的参数？

本章小结

学习本章要求掌握图像退化模型、图像退化原因、图像复原的各种方法以及它们之间的相互关系。本章重点要求明确图像退化模型、图像复原方法的应用。本章的主要内容可以概括如下：

(1) 图像退化的原因很多，如摄像时的散焦模糊、成像系统的噪声干扰，相机与景物之间相对运动形成的运动模糊、图像的几何失真等。对图像的复原方法有一定的针对性，根据不同的失真机理选择不同的复原方法往往是较好复原的关键。

(2) 图像退化的基本模型有连续型和离散型。连续型模型便于分析光学系统形成的点扩散函数 (PSF)。离散型模型便于用线性代数和数字信号处理理论与方法建立图像退化模型, 便于计算机求解。

(3) 线性复原方法分为非约束复原方法和约束复原方法。在同一种复原技术中, 不同的 PSF、噪声强度、算子的搜索范围、约束算子等对复原的结果影响都是不可低估的。

(4) 非线性复原方法采用的准则函数不能用线性代数方法简化运算。非线性复原的有效方法之一是迭代法。同态滤波是另一种非线性复原方法, 它对于乘积型结构图像的复原比较有效。

(5) 盲图像信号复原方法建立在对图像退化的先验知识未知的情况下, 通过观察多幅图像获取退化信息以复原图像。

(6) 几何畸变校正分为两个步骤。首先, 建立两个空间像素点之间的位置关系。然后, 根据失真图像映射到标准图像空间的位置对应关系, 采用插值法确定标准空间像素点的灰度值。

思考题与习题

7.1 引起图像退化的原因有哪些?

7.2 常见的图像退化模型包含哪些种类?

7.3 讨论在没有退化函数 H 的相关知识情况下, 利用处理过的高信噪比子图像作为原图像的估计图像, 采用图像观察估计 PSF 的方法。

7.4 逆滤波复原的病态解问题需要人为地对传递函数进行修正。以降低由于传递函数病态而造成的恢复不稳定性, 对逆滤波器的截止半径有何要求?

7.5 用维纳滤波的方法进行图像复原, 不同的 PSF 对复原效果有什么影响?

7.6 用约束最小二乘方滤波复原时, 不同的噪声强度、拉普拉斯算子的搜索范围和约束算子对复原效果有何影响?

7.7 盲去卷积方法中, 如何选择合适的 PSF 值?

7.8 对于几何失真校正, 尝试用 Photoshop “滤镜” 菜单、“镜头校正” 命令实现。

第8章 图像分割

➡ 内容提要

图像分割的主要目的是将目标从背景中提取出来，是为图像分析和图像理解做准备的处理过程。由于目标的特征不同、提取特征的方法不同，因而图像分割的方法繁多。本章主要介绍图像分割的基本原理和方法，包括图像的阈值分割技术、图像的边缘检测和区域检测法等。本章还给出了一些分割方法的实例。

➡ 知识要点

- 图像分割的定义及其在图像处理与分析中的地位。
- 像素间的关系：邻域和连通性。
- 阈值分割技术：全局阈值分割和自适应阈值分割技术。
- 边缘检测法：梯度算子、拉普拉斯算子、拉普拉斯-高斯算子、方向算子、坎尼算子和边缘跟踪。
- 区域检测法。
- 霍夫变换。

➡ 教学建议

- 本章教学安排 4 学时。
- 学习本章的先修知识包括集合运算、求极值、梯度算子等。
- 重点了解各种图像分割方法的特点，根据图像的特点采用合适的分割方法并使用 MATLAB 的相关函数进行图像分割。

8.1 概述

8.1.1 图像分割的目的和任务

前面几章主要针对图像的整体状况进行处理。图像经过处理后,得到的是改善后的另一幅图像,或更利于有效应用的另一幅图像。而在对图像进行分析的另一些场合,人们可能对图像中的某些局部或特征感兴趣,其输出不一定是一幅完整的图像。这些部分常被称为目标或对象(object),处于感兴趣的区域(Region of Interest, ROI)。在图像分析中,输出的结果是对图像的描述、分类或其他某种结论,而不再像常规图像处理那样输出另一幅图像。由于这些被分割的区域在某些特性上相近,因而,图像分割常用于模式识别与图像理解以及图像压缩与编码等应用场合。

由此可见,图像分割是图像处理到图像分析的关键步骤,是计算机视觉的基础,是图像理解的重要组成部分。如图8.1所示,目标为飞机,分割后的图像中,飞机的外形特征比原图像更加明显。

一个典型的图像分析和理解系统分为图像采集、图像预处理、图像分割、特征提取、图像识别与分类、结论等部分。图像分割根据目标与背景的先验知识将图像表示为物理上有意义的连通区域的集合,即对图像中的

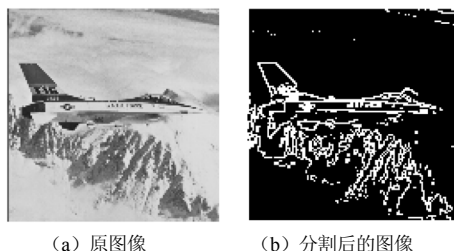


图8.1 目标为飞机的图像

目标、背景进行标记、定位,然后将目标从背景或其他伪目标中分离出来。图像分割以后,通常需要对分割的区域进行表示(representation)和描述(description),以便计算机进一步处理。区域既可以用外部边界特征表示,也可以由组成区域的内部像素表示。在区域表示的基础上,区域描述用一组数量或符号来表征图像中被描述目标的某些特征,是对图像中各组成部分之间关系或对各部分性质的描述。

图像分割的困难在于图像数据的模糊和噪声干扰。实际图像中,景物情况各异,需要根据实际情况选择适合的方法。分割结果的好坏或者正确与否,目前还没有一个统一的评价判断准则,大多从分割的视觉效果和实际的应用场景来判断。图像分割质量评价通过对图像分割算法的性能研究来筛选出更优的分割算法以达到优化分割的目的。为使评价方法实用准确,对评价方法提出了以下基本要求:

① 具有通用性,即评价方法能够适用于不同的分割算法及各种应用领域;② 采用定量的和客观的性能评价准则;③ 选取通用的图像作为参照进行测试以使各评价结果具有可比性。

8.1.2 图像分割的集合定义

常用集合来定义图像分割(image segment)。令集合 R 代表整个图像区域,对 R 的图像分割可以视为将 R 分成 N 个满足以下条件的非空子集 R_1, R_2, \dots, R_N :

- (1) $\bigcup_{i=1}^N R_i = R$;
- (2) 对于所有的 i 和 j , $i \neq j$, 有 $R_i \cap R_j = \emptyset$;
- (3) 对于 $i=1, 2, \dots, N$, 有 $P(R_i) = \text{TRUE}$;
- (4) 对于 $i \neq j$, 有 $P(R_i \cup R_j) = \text{FALSE}$;
- (5) 对于 $i=1, 2, \dots, N$, R_i 是连通的区域。

其中, $P(R_i)$ 是对所有在集合 R_i 中的元素的逻辑谓词, \emptyset 代表空集。条件(1)表示分割的所有子

区域的并集就是原来的图像。这一点非常重要，因为它保证了图像处理中的每个像素都被处理的充分条件。条件（2）表明分割结果中各个区域是互不重叠的。条件（3）指出在分割结果中，每个区域都有独特的特性。条件（4）表示在分割结果中，不同的子区域具有不同的特性，它们没有公共的特性。条件（5）则要求分割结果中同一个子区域内的像素应当是连通的，即同一个子区域内的任意两个像素在该子区域内互相连通。

这些条件对分割有一定的指导作用。但是，实际的图像处理和分析都是面向某种特定的应用的，所以条件中的各种关系也是需要和实际需求结合起来设定的。迄今为止，还没有一种通用的方法可以把人类的要求完全转换成图像分割中的各种条件关系，条件表达式往往都是近似的。

8.1.3 图像分割的分类

图像是千差万别的，图像的分割方法也是丰富多彩的。图像分割除依照图像自身的特点进行处理外，数学和信号处理领域新的理论和方法，往往被人们引入到分割的算法中。因而出现了基于模糊数学的图像分割、数学形态学的图像分割、基于神经网络的分割、基于遗传算法等优化理论的分割等。

根据分割方法的不同，通常有两种分类方法：

（1）根据图像的两种特性进行分割：一种是根据各个像素点的灰度不连续性进行分割；另一种是根据同一区域具有相似的灰度进行分割。

（2）根据分割的处理策略不同进行分割：一种是并行算法，所有的判断和决策可以独立进行；另一种是串行算法，后期的处理依赖前期的运算结果。后者运算时间较长，但抗干扰能力较强。

表 8.1 列举了常见的图像分割算法。

表 8.1 常见的图像分割算法

分 类	边界（不连续性）	区域（相似性）
并行处理	并行边界类（边缘检测等）	并行区域类（阈值分割、聚类等）
串行处理	串行边界类（边缘跟踪等）	串行区域类（区域生长、分裂合并等）

根据应用目的不同，分为粗分割和细分割：对于模式识别应用，一个物体对象内部的细节与颜色或灰度渐变应被忽略，而且一个物体对象只应被表示为一个或少数几个分割区域，即粗分割；而对于基于区域或对象的图像压缩与编码，其分割的目的是为了得到颜色或灰度信息一致的区域，以利于高效的区域编码。若同一区域内含有大量变化细节，则难以编码，图像需要细分割，即需要捕捉图像的细微变化。

根据分割对象的属性，可分为灰度图像分割和彩色图像分割。

根据分割对象的状态，可分为静态图像分割和动态图像分割。

根据分割对象的应用领域，可分为遥感图像分割、交通图像分割、医学图像分割、工业图像分割、军事图像分割等。

8.2 像素的邻域和连通性

像素的邻域通常有两种定义方法。

（1）4 邻域：对一个坐标为 (x, y) 的像素 p ，它可以有两个水平和两个垂直的近邻像素。互为 4 邻域的像素又称为 4 连通的。

（2）8 邻域：取像素 p 四周的 8 个点作为相连接的邻域点。互为 8 邻域的像素又称为 8 连通的。

```
0 0 0 0 0
0 1 1 0 0
0 1 0 1 0
0 1 1 1 0
0 0 0 0 0
```

值得注意的是，目标和背景必须取不同的连通性定义，否则会引起矛

图 8.2 目标和背景连通性

盾。如图 8.2 所示, 1 表示目标, 0 表示背景, 则目标定义为 8 连通, 背景定义为 4 连通, 从而灰度为 1 的 5 个点可以连接成封闭环, 中间的 0 就变成了洞, 它和环外的区域不相连通。

根据连通性 (connectivity), 可以定义图像的特征点和线。

(1) 边界点 (boundary point): 如果目标点集 S 中的点 p , 有邻点在 S 的补集 \bar{S} 中, 那么 p 便是 S 的边界点。这种点的集合, 便是 S 的边界, 记为 S' 。

(2) S 的内部 (interior) 和内点 (interior point): 目标点集 S 和边界点集 S' 之差 $S - S'$ 称为 S 的内部, 处于 S 内部的点称为 S 的内点。

(3) 孤点 (isolated point): 没有邻接点的点。

(4) 封闭曲线 (closed curve): 如果连通域 S 中的所有点都有两个邻点, 则称此连通域为封闭曲线。

各种特征点、线的实例见图 8.3。

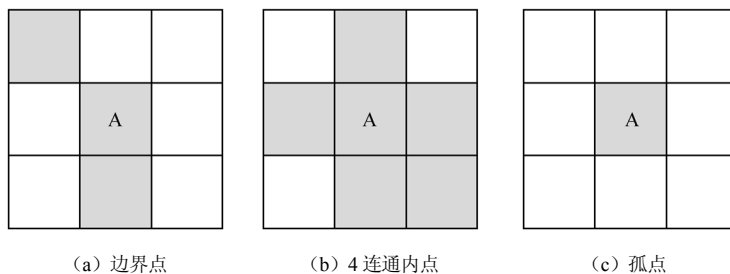


图 8.3 各种特征点、线的实例

【例 8.1】 根据 4 连通或 8 连通准则在二值图像中判断目标。

【解】 应用 MATLAB 提供的 bwlablel 函数可以根据 4 连通或 8 连通准则, 在给定的二值图像矩阵 BW 中寻找目标。MATLAB 程序如下:

```
BW = [1 1 1 0 0 0 0 0;
      1 1 1 0 1 1 0 0;
      1 1 1 0 1 1 0 0;
      1 1 1 0 0 0 1 0;
      1 1 1 0 0 0 1 0;
      1 1 1 0 0 0 1 0;
      1 1 1 0 0 1 1 0;
      1 1 1 0 0 0 0 0]; % 给定的二值图像矩阵
L4 = bwlablel(BW,4)      % 根据 4 连通准则判定目标
L8 = bwlablel(BW,8)      % 根据 8 连通准则判定目标
```

根据 4 连通准则, 得到的目标是 3 个, 结果如下:

```
L4 =
      1  1  1  0  0  0  0  0
      1  1  1  0  2  2  0  0
      1  1  1  0  2  2  0  0
      1  1  1  0  0  0  3  0
      1  1  1  0  0  0  3  0
      1  1  1  0  0  0  3  0
      1  1  1  0  0  3  3  0
      1  1  1  0  0  0  0  0
```

根据 8 连通准则, 得到目标是 2 个, 结果如下:

```
L8 =
    1    1    1    0    0    0    0    0
    1    1    1    0    2    2    0    0
    1    1    1    0    2    2    0    0
    1    1    1    0    0    0    2    0
    1    1    1    0    0    0    2    0
    1    1    1    0    0    0    2    0
    1    1    1    0    0    2    2    0
    1    1    1    0    0    0    0    0
```

8.3 图像的阈值分割技术

8.3.1 基本原理

图像的阈值分割技术属于一类并行区域分割算法。若图像中目标和背景具有明显不同的灰度集合, 且两个灰度集合可用一个灰度级阈值 T 进行分割, 那么就可以用阈值分割灰度级的方法在图像中分割出目标区域与背景区域。在物体与背景有较强对比度的图像中, 应用此种方法特别有效, 例如印刷文字图像。

设图像为 $f(x, y)$, 其灰度集范围是 $[Z_1, Z_K]$, 在 Z_1 和 Z_K 之间选择一个合适的灰度阈值 T , 则图像分割方法可由下式描述

$$g(x, y) = \begin{cases} Z_E, & f(x, y) \geq T \\ Z_B, & f(x, y) < T \end{cases} \quad (8.1)$$

这样得到的 $g(x, y)$ 是一幅目标灰度为 Z_E 、背景灰度为 Z_B 的二值图像。

【例 8.2】 给出利用阈值分割图像 rice.tif 的实例。

【解】 该图像的目标 (米粒) 和背景 (偏暗) 比较分明, 用阈值 ($T = 110$) 分割算法将达到较好的分割效果。程序如下:

```
I = imread('rice.tif');
subplot(131); imshow(I);
subplot(132); imhist(I);
T = 110;
S = size(I)
[maxI, maxP] = max(I(:,:)) %maxI 为图像 I 的最大灰度值, maxP 为其位置
[minI, minP] = min(I(:,:)) %minI 为图像 I 的最小灰度值, minP 为其位置
for i = 1:S(1)
    for j = 1:S(2)
        if I(i,j) >= T I(i,j) = 255;
        else I(i,j) = 0;
    end
end
end
subplot(133); imshow(I);
```

运行结果如图 8.4 所示, 图 8.4 (a) 是待分割的原图像, 图 8.4 (b) 是对应的直方图, 图 8.4 (c) 是选择分割阈值为 110 的结果图。程序运行的数据如下:

```

S = 256  256 (图像大小为 256×256)
maxI = 255 (即  $z_K = 255$ )
maxP = 24908
minI = 40 (即  $z_1 = 40$ )
minP = 234

```

改变阈值, 可以发现过低或过高的阈值都会影响分割的效果。可见, 阈值的选择是该算法的关键因素。

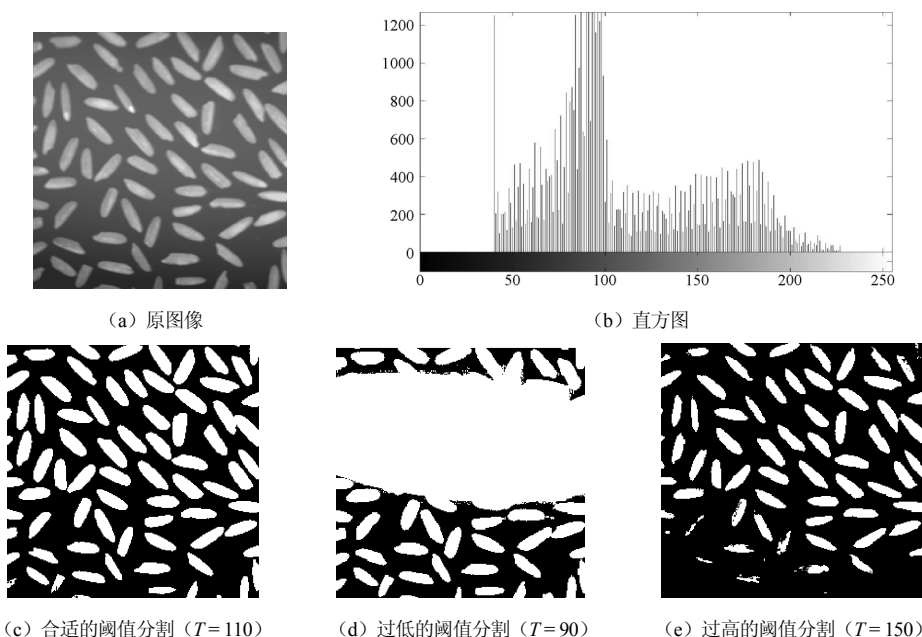


图 8.4 阈值分割

阈值的选取方法一般可以分为全局阈值法和局部阈值法两类。如果分割过程中对图像上每个像素所使用的阈值相等, 则为全局阈值方法; 如果每个像素所使用的阈值不同, 则为局部阈值方法。

8.3.2 全局阈值分割

全局阈值 (global thresholding) 是最简单的图像分割方法。根据不同的目标, 选用最佳的阈值。常用以下几种方法确定分割的最佳阈值。

(1) 实验法: 如果分割之前就知道图像的一些特征, 那么阈值确定就比较简单, 只要用不同的阈值进行测试, 即可检查该阈值是否适合图像的已知特征。这种方法需要知道图像的某些特征, 但有时这些特征是不可预知的。

(2) 直方图法: 先作出图像的灰度直方图。如图 8.5 所示, 若其直方图呈双峰且有明显的谷底, 则可以将谷底点所对应的灰度值作为阈值 T , 然后根据该阈值进行分割, 就可以将目标从图像中分割出来。这种方法适用于目标和背景的灰度差较大且直方图有明显谷底的情况。例 8.2 的分割策略就属于这种情况。

(3) 最小误差的方法: 假设背景的概率密度为 $p_1(z)$, 感兴趣目标的概率密度为 $p_2(z)$ 。设背景像素数占图像总像素数的百分比为 θ , 前景 (即目标) 的像素数百分比为 $(1-\theta)$, 如图 8.6 所示, 则混合概率密度为

$$p(z) = \theta p_1(z) + (1-\theta) p_2(z) \quad (8.2)$$

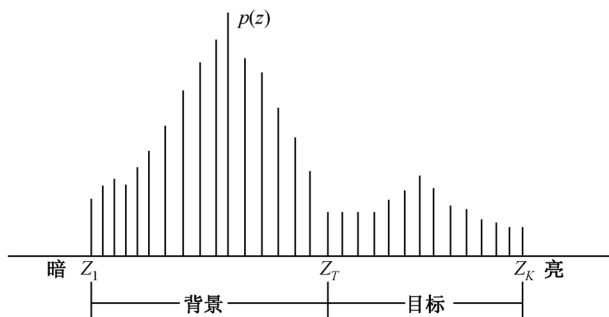


图 8.5 直方图阈值分割示意图

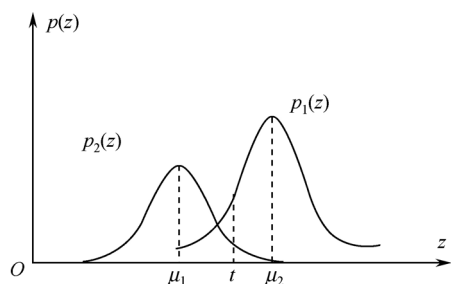


图 8.6 灰度级分布

当选定阈值为 t 时，目标像素错分为背景像素的概率为

$$E_2(t) = \int_t^{\infty} p_2(z) dz$$

把背景像素错分为目标像素的概率为

$$E_1(t) = \int_{-\infty}^t p_1(z) dz$$

总错误概率为

$$E(t) = \theta E_1(t) + (1 - \theta) E_2(t) \quad (8.3)$$

为了使这个误差最小，可令 $\frac{\partial E(t)}{\partial t} = 0$ ，则有

$$-\theta p_1(t) + (1 - \theta) p_2(t) = 0 \quad (8.4)$$

用上述方法求得的阈值为最佳阈值。假设背景与前景的灰度分布都是正态分布的，背景的均值和方差分别为 μ_1 和 σ_1^2 ，前景的均值和方差分别为 μ_2 和 σ_2^2 ，即

$$p_1(z) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(z-\mu_1)^2}{2\sigma_1^2}} \quad (8.5)$$

$$p_2(z) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(z-\mu_2)^2}{2\sigma_2^2}} \quad (8.6)$$

将式 (8.5) 和式 (8.6) 代入式 (8.4) 中，并通过对数运算可以得出

$$\ln \frac{\theta \sigma_2}{(1 - \theta) \sigma_1} - \frac{(t - \mu_1)^2}{2\sigma_1^2} = \frac{-(t - \mu_2)^2}{2\sigma_2^2} \quad (8.7)$$

当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时，有

$$t = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln \frac{1 - \theta}{\theta} \quad (8.8)$$

如果先验概率已知，例如 $\theta = \frac{1}{2}$ ，则有

$$t = \frac{\mu_1 + \mu_2}{2} \quad (8.9)$$

这表示如果目标和背景是正态分布的，最佳阈值可以按式(8.8)或式(8.9)来求。若 $p_1(z)$ 和 $p_2(z)$ 不是正态分布，则可根据式(8.4)来计算阈值 t 。

(4) 最大类间方差法

该法又称为大津算法(Otsu's method)。算法选择阈值使得目标和背景类间方差(inter-class variance)最大化。目标(设为 C_1 类)和背景(设为 C_2 类)之间的类间方差越大，说明构成图像的两部分的差别越大。大津指出最小化类内方差(intra-class variance)与最大化类间方差是等价的。当部分目标错分为背景或部分背景错分为目标都会导致两部分差别变小。该算法因其简单、快速、分割精确、适用范围广而成为广泛采用的一种图像阈值分割方法。

设图像的像素数为 N ，灰度范围为 $[0, L-1]$ ，灰度级为 $i \in [0, L-1]$ 的像素数为 n_i ，概率为

$$p_i = n_i / N \quad (8.10)$$

且 $\sum_{i=0}^{L-1} p_i = 1$ ， $\theta = \sum_{i=0}^t p_i$ 。 C_1 类的灰度值 $\in [0, t]$ ， C_2 类的灰度值 $\in [t+1, L-1]$ ，则图像的总均值、 C_1 类和 C_2 类的均值分别为 μ 、 μ_1 和 μ_2

$$\mu = \sum_{i=0}^{L-1} ip_i, \quad \mu_1 = \sum_{i=0}^t ip_i / \theta, \quad \mu_2 = \sum_{i=t+1}^{L-1} ip_i / (1-\theta) \quad (8.11)$$

由此可得

$$\mu = \theta\mu_1 + (1-\theta)\mu_2 \quad (8.12)$$

类间方差定义为

$$\sigma^2 = \theta(\mu - \mu_1)^2 + (1-\theta)(\mu_2 - \mu)^2 \quad (8.13)$$

将式(8.12)代入式(8.13)，得

$$\sigma^2 = \theta(1-\theta)(\mu_2 - \mu_1)^2 \quad (8.14)$$

上述 θ 、 μ_1 和 μ_2 均为 t 的因变量，采用遍历 t 的方法可得到使类间方差最大的阈值 t ，即为所求。

IPT 提供了 `graythresh` 函数利用大津算法(Otsu's method)计算全局灰度阈值。调用格式为

$$\text{LEVEL} = \text{graythresh}(I)$$

其中， I 为灰度图像，`level` 为返回的归一化全局阈值(即 `level` $\in [0, 1]$)。

进一步，可调用 `im2bw` 函数按阈值 `level` 将灰度图像 I 转换成二值图像 BW 。调用格式为

$$BW = \text{im2bw}(I, \text{level})$$

【例 8.3】原始图像 `cameraman.tif` 如图 8.7(a) 所示，使用函数 `graythresh` 对其进行全局阈值分割。求出最佳阈值，显示分割效果图。

【解】程序代码如下：

```
I = imread('cameraman.tif');
figure(1); imshow(I);
level = graythresh(I)
BW = im2bw(I, level);
figure(2); imshow(BW)
```

程序运行结果如图 8.7 所示。

level = 0.3451

对 level 去归一化, 对于 8 位灰度图像, 有 $L=256$, 可求实际阈值为 $(L-1) \times \text{level} = 88$ 。实际分割效果是较为理想的。



图 8.7 大津算法阈值分割效果图

以图 6.9 所示的显微图像为例。由灰度直方图可见有两个波峰, 双峰之间谷底处的灰度值作为阈值进行灰度分割, 取 4 个较为合适的分割点: $T=80$ 、 $T=100$ 、 $T=120$ 、 $T=140$, 阈值分割效果如图 8.8 所示, $T=80$ 时提取 SBS 颗粒较多, $T=140$ 时, 提取 SBS 颗粒较少。

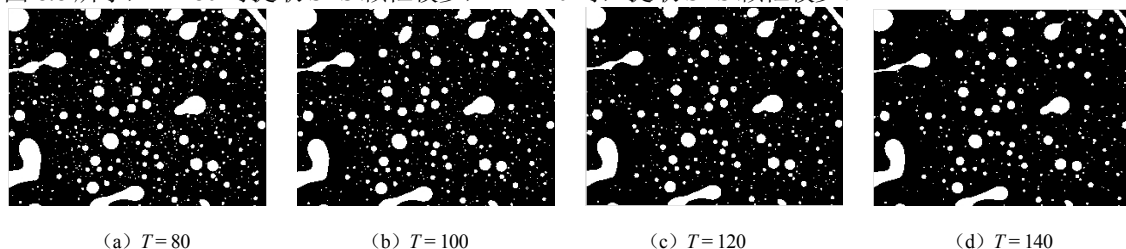


图8.8 显微图像阈值分割

8.3.3 局部阈值分割

局部阈值法常用于照度不均或灰度连续变化的图像分割, 又称为自适应阈值分割法。当照明不均匀、有突发噪声或者背景灰度变化比较大的时候, 单一的阈值不能兼顾图像各个像素的实际情况。这时, 可以对图像进行分块处理, 对每一块分别选定一个阈值进行分割, 这种与坐标相关的阈值分割法称为动态阈值方法, 也称为自适应阈值的方法。

这类算法的时间复杂度和空间复杂度比较大, 但是抗噪声的能力比较强, 对采用全局阈值不容易分割的图像有较好的效果。这种方法的关键问题是如何将图像进行细分和如何为得到的子图像估计阈值。由于用于每个像素的阈值取决于像素在子图像中的位置, 这类阈值处理是自适应的。

8.4 图像的边缘检测

8.4.1 边缘检测的基本原理

上述的阈值法是基于像素的分割方法。实验表明, 人们对图像中边缘的识别不是通过设置阈值来分割的, 目标的边缘一般表现为灰度 (对彩色图像还包括色度) 的特变。对于人类的视觉感知, 图像边缘对理解图像内容起到关键作用。如图 8.9 所示的在灰度渐变的图像中无法区分其灰度变化的边界, 但如果边界灰度有突变, 则可以区分两个灰度不同的区域。这是基于灰度不连续性进行的分割方法。



图 8.9 几种常见的边缘

边缘的含义体现在灰度的突变上。在第 6 章中, 我们讨论了如何使用差分、梯度、拉普拉斯算子及各种高通滤波处理方法来对图像边缘进行增强。事实上, 只要再进行一次阈值处理, 便可以将边缘增强的方法用于边缘检测。但是需要注意的是, 对边缘处理的目的是已经不是对整幅图像的边缘进行加强, 而是根据边缘来进行图像分割。边缘检测要按照图像的内容和应用的要求进行, 可以先对图像做预处理, 使边缘突出, 然后选择合适的阈值进行分割。

MATLAB 边缘检测的函数为

$BW = \text{edge}(I, \text{'edge-finding methods'}, \text{THRESH}, \text{DIRECTION}/\text{SIGMA})$

或

$BW = \text{edge}(I, \text{'edge-finding methods'}, \text{THRESH}, \text{SIGMA})$

其中, I 为灰度图像, BW 为返回的相同大小的二值图像。若检测到边沿则返回 1, 否则返回 0。

edge 函数支持 6 种不同的边缘检测方法。'sobel' 指定 Sobel 方法, 'prewitt sobel' 指定 Prewitt 方法, 'roberts' 指定 Roberts 方法, 'log' 指定 LoG 方法 (高斯-拉普拉斯算子), 'zerocross' 指定 (零交叉法), 'canny' 指定坎尼方法 (SIGMA 为高斯滤波器的标准差)。

THRESH 指定分割方法的阈值。

DIRECTION 指定分割方法的方向, 可以指定 'horizontal' (水平) 或 'vertical' (垂直) 边缘, 或 'both' (双向) 边缘 (默认值)。

如果 THRESH、DIRECTION、SIGMA 默认, 则由 MATLAB 指定相应的数值。

8.4.2 梯度算子

第 6 章介绍了梯度对应于一阶导数, 相应的梯度算子就对应于一阶导数算子。这里与第 6 章的不同之处在于通过算子检测后, 还需要做二值处理, 从而找到边界点。用于边缘检测的梯度算子主要有 Roberts 算子、Prewitt 算子、Sobel 算子。这三种检测算子中, Roberts 算子定位精度较高, 但也易丢失部分边缘, 抗噪声能力差, 适用于低噪声、陡峭边缘的场合。Prewitt 算子、Sobel 算子首先对图像做平滑处理, 因此具有一定的抑制噪声的能力, 但不能排除检测结果中的虚假边缘, 易出现多像素宽度。

【例 8.4】 给出利用梯度算子对图像 blood1.tif 进行边缘检测的实例。

【解】 图 8.9 (b) 至图 8.9 (d) 分别给出了利用这三个算子进行边缘检测的不同效果。

完成上述检测的实现程序如下:

```
I = imread('blood1.tif');
imshow(I);
BW1 = edge(I, 'roberts'); %进行 Roberts 算子边缘检测, 门限值采用默认值
BW2 = edge(I, 'prewitt'); %进行 Prewitt 算子边缘检测, 门限值采用默认值
BW3 = edge(I, 'sobel'); %进行 Sobel 算子边缘检测, 门限值采用默认值
BW4 = edge(I, 'log'); %进行 LoG 算子边缘检测, 门限值采用默认值
BW5 = edge(I, 'canny'); %进行 canny 算子边缘检测, 门限值采用默认值
figure, imshow(BW1, []);
figure, imshow(BW2, []);
figure, imshow(BW3, []);
figure, imshow(BW4, []);
figure, imshow(BW5, []);
```

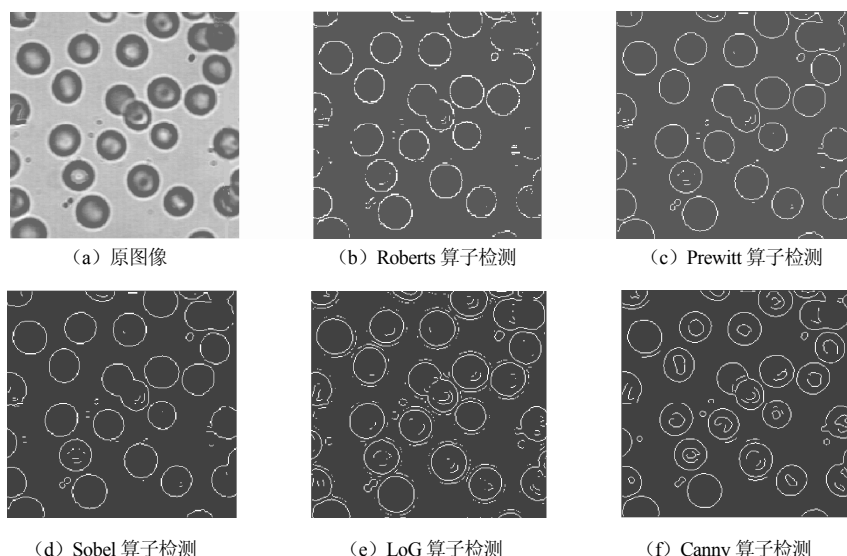



图 8.9 不同微分算子的边缘检测效果

8.4.3 拉普拉斯算子

以上梯度算子均是一阶导数算子，而拉普拉斯算子是二阶导数算子。下面讨论的二阶算子也可以检测边缘，其基本原理是检测阶梯状边缘时，需要将算子与图像进行卷积并确定过零点。图 8.10 (a) 表示模糊边缘斜坡而不是突变信号，则边缘的位置选在梯度的极大值处，即二阶导数算子的零交叉处。图 8.10 (b) 表示若边缘是一个斜坡，则边缘的梯度值相等，拉普拉斯算子为零，此时，确定边缘位置在拉普拉斯算子值的两个峰值中间。

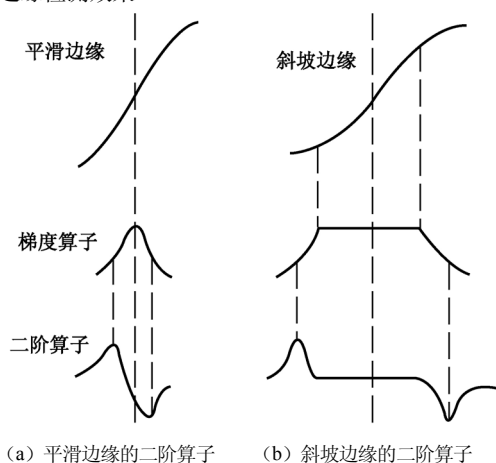


图 8.10 二阶导数算子确定边缘位置

拉普拉斯算子对图像中阶跃型边缘定位准确，但对噪声具有极高的敏感性，可能丢失部分边缘的方向信息，造成不连续的检测边缘，所以在边缘检测中用得比较少。如果先对图像做平滑降噪处理，则可以明显地降低对噪声的敏感性。拉普拉斯-高斯算子正是基于这种思想提出来的。

【例 8.5】以 Lena 图像为例，给出 Robert 算子、Sobel 算子和 Laplace 算子的边缘检测结果，如图 8.11 所示。

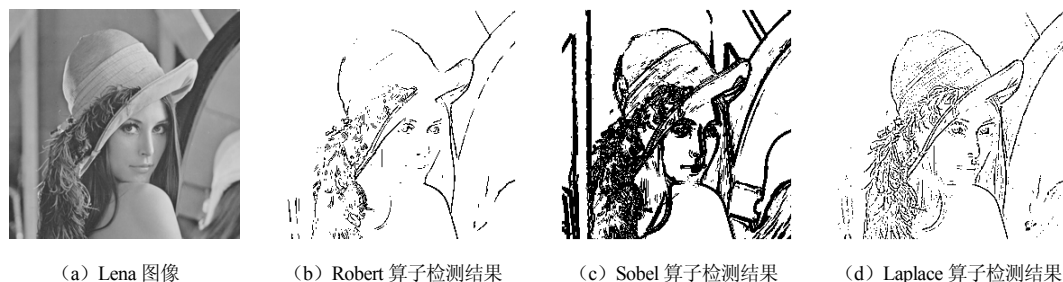


图 8.11 各种算子的检测结果

8.4.4 拉普拉斯-高斯算子

拉普拉斯-高斯算子（Laplacian of Gaussian, LoG），也称 Marr 算子。其基本思想是，先用高斯函数 $g(x, y)$ 对图像 $f(x, y)$ 进行滤波，再对滤波后的图像进行拉普拉斯运算（ ∇^2 ），结果为零的位置即为边缘点的位置。滤波提高了抗噪声的能力，但同时可能使原本比较尖锐的边缘平滑了，甚至无法检测到。

图 8.9（e）所示是应用 LoG 算子对图 8.9（a）进行边缘检测的结果。

8.4.5 坎尼边缘检测算子

坎尼边缘检测算子的梯度是用高斯滤波器的导数计算的，边缘出现在梯度的局部极大值处。坎尼的主要工作是推导了最优边缘检测算子。考核边缘检测算子的指标是：低误判率；高定位精度；单像素边缘，抑制虚假边缘。坎尼算子不易受噪声干扰，但边缘的连续性不如 LoG 算子。

坎尼算法的步骤如下。

（1）利用高斯函数平滑图像。设高斯函数为

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (8.15)$$

计算二维卷积 $\nabla G(x, y) * f(x, y)$ 实现图像平滑。

（2）计算滤波后的边缘强度和方向，通过阈值来检测边缘。将 $\nabla G(x, y)$ 的二维卷积模板分解为两个一维滤波器

$$\frac{\partial G(x, y)}{\partial x} = kxe^{-\frac{x^2}{2\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} = h_1(x)h_2(y) \quad (8.16)$$

$$\frac{\partial G(x, y)}{\partial y} = kye^{-\frac{y^2}{2\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} = h_1(y)h_2(x) \quad (8.17)$$

$$\text{式中} \quad h_1(x) = \sqrt{k}xe^{-\frac{x^2}{2\sigma^2}}; \quad h_2(y) = \sqrt{k}e^{-\frac{y^2}{2\sigma^2}} \quad (8.18)$$

$$h_1(y) = \sqrt{k}ye^{-\frac{y^2}{2\sigma^2}}; \quad h_2(x) = \sqrt{k}e^{-\frac{x^2}{2\sigma^2}} \quad (8.19)$$

$$\text{可见} \quad h_1(x) = xh_2(x); \quad h_1(y) = yh_2(y) \quad (8.20)$$

然后把这两个模板分别与 $f(x, y)$ 进行卷积，得到

$$E_x = \frac{\partial G(x, y)}{\partial x} * f; \quad E_y = \frac{\partial G(x, y)}{\partial y} * f \quad (8.21)$$

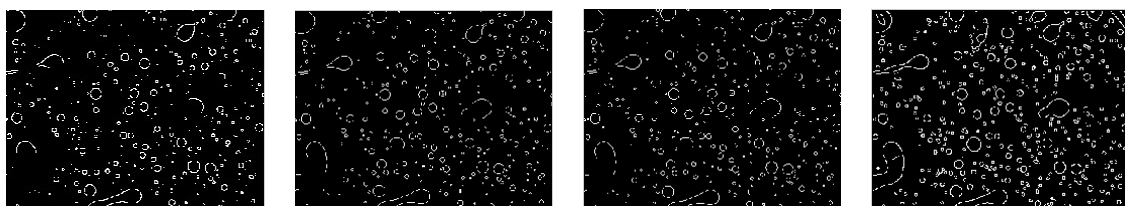
$$\text{令} \quad A(x, y) = \sqrt{E_x^2 + E_y^2}; \quad a(x, y) = \arctan \left[\frac{E_y(x, y)}{E_x(x, y)} \right] \quad (8.22)$$

则 A 反映边缘强度， a 为垂直于边缘的方向。

（3）判断一个像素是否为边缘点的条件为：像素 (x, y) 的边缘强度大于沿梯度方向的两个相邻像素的边缘强度；与该像素梯度方向上相邻两点的方向差小于 45° ；以该像素为中心的 3×3 邻域中的边缘强度的极大值小于某个阈值。

图 8.8（f）所示是应用坎尼算子对图 8.8（a）进行边缘检测的结果。可见坎尼算子的检测比较优越，它可以减少小模板检测中的边缘中断，有利于得到较完整的边缘。

以显微图像为例考察不同边缘检测算子的性能。如图 8.12 所示, 边缘检测可以更清楚地显示出颗粒的边缘轮廓, 便于观察。采用 4 种算子进行边缘检测, Canny 算子检测效果最好、其他与例 8.4 一致。



(a) Roberts 算子

(b) Prewitt 算子

(c) Sobel 算子

(d) Canny 算子

图8.12 显微图像边缘检测

8.4.6 方向算子

与梯度算子不同的是, 方向算子是利用一组模板对图像中的同一像素求卷积, 然后选取其中最大的值作为边缘强度, 而将与之对应的方向作为边缘方向。它相对于梯度算子的优点是不仅仅只考虑水平和垂直方向, 还可以检测其他方向上的边缘, 但计算量将大大增加。

常用的有 8 方向 Kirsch (3×3) 模板, 如图 8.13 所示, 方向间的夹角为 45° 。

<table><tr><td>-5</td><td>3</td><td>3</td></tr><tr><td>-5</td><td>0</td><td>3</td></tr><tr><td>-5</td><td>3</td><td>3</td></tr></table>	-5	3	3	-5	0	3	-5	3	3	<table><tr><td>3</td><td>3</td><td>3</td></tr><tr><td>-5</td><td>0</td><td>3</td></tr><tr><td>-5</td><td>-5</td><td>3</td></tr></table>	3	3	3	-5	0	3	-5	-5	3	<table><tr><td>3</td><td>3</td><td>3</td></tr><tr><td>3</td><td>0</td><td>3</td></tr><tr><td>-5</td><td>-5</td><td>-5</td></tr></table>	3	3	3	3	0	3	-5	-5	-5	<table><tr><td>3</td><td>3</td><td>3</td></tr><tr><td>3</td><td>0</td><td>-5</td></tr><tr><td>3</td><td>-5</td><td>-5</td></tr></table>	3	3	3	3	0	-5	3	-5	-5
-5	3	3																																					
-5	0	3																																					
-5	3	3																																					
3	3	3																																					
-5	0	3																																					
-5	-5	3																																					
3	3	3																																					
3	0	3																																					
-5	-5	-5																																					
3	3	3																																					
3	0	-5																																					
3	-5	-5																																					
<table><tr><td>3</td><td>3</td><td>-5</td></tr><tr><td>3</td><td>0</td><td>-5</td></tr><tr><td>3</td><td>3</td><td>-5</td></tr></table>	3	3	-5	3	0	-5	3	3	-5	<table><tr><td>3</td><td>-5</td><td>-5</td></tr><tr><td>3</td><td>0</td><td>-5</td></tr><tr><td>3</td><td>3</td><td>3</td></tr></table>	3	-5	-5	3	0	-5	3	3	3	<table><tr><td>-5</td><td>-5</td><td>-5</td></tr><tr><td>3</td><td>0</td><td>3</td></tr><tr><td>3</td><td>3</td><td>3</td></tr></table>	-5	-5	-5	3	0	3	3	3	3	<table><tr><td>-5</td><td>-5</td><td>3</td></tr><tr><td>-5</td><td>0</td><td>3</td></tr><tr><td>3</td><td>3</td><td>3</td></tr></table>	-5	-5	3	-5	0	3	3	3	3
3	3	-5																																					
3	0	-5																																					
3	3	-5																																					
3	-5	-5																																					
3	0	-5																																					
3	3	3																																					
-5	-5	-5																																					
3	0	3																																					
3	3	3																																					
-5	-5	3																																					
-5	0	3																																					
3	3	3																																					

图 8.13 3×3 Kirsch 算子的 8 方向模板

8.4.7 边缘跟踪

在一些应用场合, 仅仅得到边缘点是不够的。由于噪声和光照不均等因素会使得原本连续的边缘出现间断现象, 因此, 在使用边缘检测算法后, 有必要采用边缘跟踪方法将间断的边缘转换成有意义的边缘信息。

基本的跟踪方法是从图像的一个边缘点出发, 根据某种判别准则, 寻找下一个边缘点, 以此形成目标的边界。起始点的选择十分重要, 起始点不同可能导致不同的跟踪结果。同样, 跟踪的终点由搜索的终止条件决定。

光栅扫描跟踪法是一种简单的利用局部信息、通过扫描的方式将边缘点连接起来的方法。

图 8.14 (a) 是一幅含有 3 条曲线的模糊图像。假设在任一点上, 曲线斜率都不超过 90° , 现在要从该图中检测出这些曲线。跟踪的具体步骤如下:

(1) 确定一个比较高的阈值 d , 把高于阈值的像素作为检出点。称该阈值为“检测阈值”, 在本例中选 $d = 7$ 。

(2) 用检测阈值 d 逐行对像素进行检测, 凡超过或等于 d 的点都接受为检出点。本例检测结果如图 8.14 (b) 所示。

(3) 选取一个比较低的阈值 T 作为跟踪阈值, 该阈值可以根据不同准则来选择。例如, 本例中是根据相邻对象点的灰度差所能允许的最大值来选择的, 取 4 作为跟踪阈值。

(4) 确定跟踪邻域。本例中取像素 (i, j) 的下一行像素 $(i+1, j-1)$ 、 $(i+1, j)$ 和 $(i+1, j+1)$ 为跟踪邻域。

(5) 从第一行开始进行检测, 找出第一行中由 d 确定的检出点作为对象点, 扫描下一行像素, 凡位于上一行已检测出来的对象点的跟踪邻域的像素, 其灰度差小于或等于跟踪阈值 t 的, 都接受为对象点, 反之去除。

(6) 对于已检测出的某一对象点, 如果在下一行跟踪邻域中, 没有任何一个像素被接受为对象点, 那么这一条曲线的跟踪可结束。如果同时有两个甚至 3 个邻域点均被接受为对象点, 则说明曲线发生分支, 跟踪将对各分支同时进行。如果若干分支曲线合并成一条曲线, 则跟踪可集中于一条曲线上进行。一条曲线跟踪结束后, 采用类似上述的步骤从第一行的其他检出点开始下一条曲线的跟踪。

(7) 对于未被接受为对象点的检出点, 再次用上述方法进行检测, 并以检出的点为起始点, 重新使用跟踪阈值程序, 以检测出不是从第一行开始的其他曲线。

(8) 当扫描完最后一行时, 跟踪便可结束。本例的跟踪结果如图 8.14 (c) 所示。

由结果可以看出, 本例原图像中存在着 3 条曲线, 两条从顶端开始, 一条从中间开始。然而, 如果不用跟踪法, 只用一种阈值 d 检测就不能得到满意的结果。

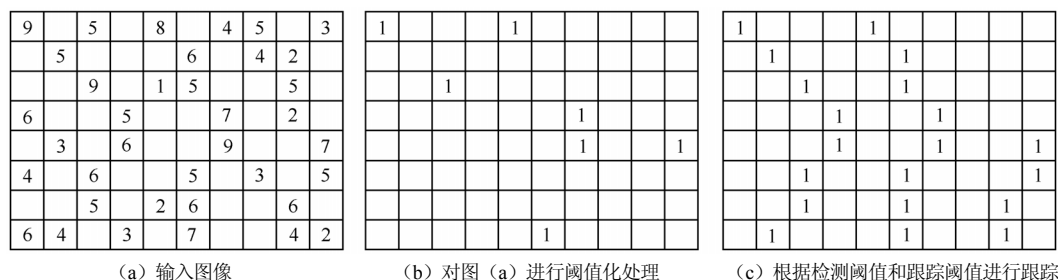


图 8.14 光栅扫描跟踪

光栅扫描跟踪和扫描方向有关, 因此最好沿其他方向再跟踪一次, 如逆向跟踪, 然后将两种跟踪的结果综合起来得到更好的结果。另外, 若边缘和光栅扫描方向平行时效果不好, 则最好在垂直扫描方向跟踪一次, 它相当于把图像转置 90° 后再进行光栅扫描跟踪。

MATLAB 提供了追踪二值图像中单个边界的函数。

(1) 形式 1

$B = \text{bwtraceboundary}(BW, P, \text{fstep})$

该函数跟踪二值图像 BW 中的目标轮廓, 返回值 B 是一个 $Q \times Q$ 矩阵, Q 是区域边界像素的数量。 B 保存有边界像素的行、列坐标。目标由非零像素组成, 零像素构成背景。 P 是一个指定行、列坐标的二元向量, 表示目标边界上的起始跟踪点。 fstep 是一个表示初始查找方向的字符串, 以寻找目标中与 P 相连的下一个像素。例如, 字符串“N”表示 north, “NE”表示 northeast。图 8.15 说明了 fstep 的所有可能值, 中央的灰色部分为当前的跟踪点。

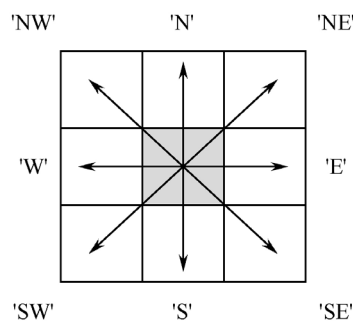


图 8.15 fstep 的所有可能值

(2) 形式2

$$B = \text{bwtraceboundary}(BW, P, \text{fstep}, \text{conn})$$

在形式1的基础上,该函数要求指定追踪边界使用的连通性 `conn`。`conn` 可以是8(默认值)和4(`fstep` 被限制为 'N'、'E'、'S'和 'W')两种数值之一。

(3) 形式3

$$B = \text{bwtraceboundary}(\dots, N, \text{dir})$$

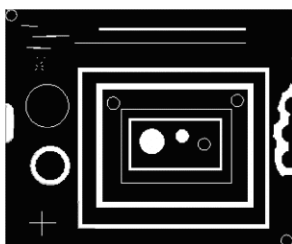
在形式2的基础上,该函数指定 `N` 为边界像素提取的最大数值, `dir` 为搜寻边界的方向。当 `N` 被设置成 `Inf` 即默认值时,则该算法识别边界上所有的像素点。`dir` 值可以取 'clockwise' (顺时针,默认值) 和 'counterclockwise' (逆时针) 两者之一。

【例 8.6】 利用 `bwtraceboundary` 函数对图像 `blobs.png` 进行边界跟踪。

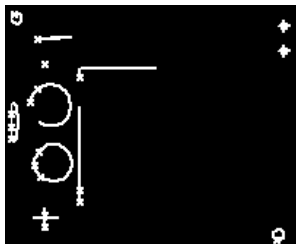
【解】 读入并显示一幅二值图像。从左上角开始,以15个像素行高为横条域,寻找第一个非零像素。使用该像素的坐标作为边界追踪的开始点。包括开始点在内,提取100个边界像素,把它们覆盖在全黑的背景图像上。使用绿色的×标记开始点,使用红色的*标记在行尾仍未发现开始点。顺时针边界跟踪的程序如下:

```
BW = imread('blobs.png');
subplot(1,2,1); imshow(BW, []);
s = size(BW);
subplot(1,2,2); background = zeros(s); imshow(background);
for row = 2:55:s(1)
    for col = 1:s(2)
        if BW(row,col) break; end
    end %寻找单一边界的第一个非零像素(跟踪开始点),坐标为[row, col]
    contour = bwtraceboundary(BW, [row, col], 'W', 8, 100, 'clockwise');
    if (~isempty(contour)) %发现跟踪起始点
        hold on; plot(col, row, 'kx', 'LineWidth', 2); %标记跟踪开始点
        hold on; plot(contour(:,2), contour(:,1), 'k', 'LineWidth', 2);
    else
        hold on; plot(col, row, 'k*', 'LineWidth', 2); %标记无跟踪开始点
    end
end
end
```

程序中 `hold on` 语句是为了将轮廓线 `contour` 覆盖在全黑的背景图像上,将 'clockwise' 改为 'counterclockwise' 可得到按逆时针边界跟踪的结果。运行结果如图 8.16 所示。



(a) 二值图像实例



(b) 顺时针边缘跟踪结果



(c) 逆时针边缘跟踪结果

图 8.16 二值图像边界跟踪

8.5 霍夫变换

霍夫变换 (Hough Transform) 是一种边界跟踪方法,它利用图像的全局特性直接检测目标轮廓。利用霍夫变换可以从图像中识别几何形状,应用很广泛,也有很多改进算法。最基本的霍夫变换是从

黑白图像中检测直线(线段)。在预先知道区域形状的条件下,利用霍夫变换可以方便地将不连续的边缘像素点连接起来得到边界曲线的逼近,其主要优点是受噪声和曲线间断的影响较小。

8.5.1 直角坐标系中的霍夫变换

霍夫变换基于点-线的对偶性(duality),即在图像空间(原空间)中同一条直线上的点对应在参数空间(变换空间)中是相交的直线。反过来,在参数空间中相交于同一点的所有直线,在图像空间中都有共线的点与之对应。

设在图像空间 XY 中,已知二值化图像中有一条直线,要求出这条直线所在的位置。由于所有过点 (x, y) 的直线一定都满足斜截式方程

$$y = px + q \quad (8.23)$$

其中 p 为斜率, q 为截距,则式(8.23)可写成

$$q = -px + y \quad (8.24)$$

式(8.24)即直角坐标中对点 (x, y) 的霍夫变换。如果将 x 和 y 视为参数,那么它也代表参数空间 PQ 中过点 (p, q) 的一条直线。

图 8.17(a) 为存在一条直线的图像空间,图 8.17(b) 是对应的参数空间。在图像空间 XY 中过点 (x_i, y_i) 的所有直线方程为 $y_i = px_i + q$, 即点 (x_i, y_i) 确定了一簇直线。它们在参数空间 PQ 中也是一条直线 $q = -px_i + y_i$ 。同理,通过点 (x_j, y_j) 的直线方程为 $y_j = px_j + q$, 它在参数空间 PQ 中是另一条直线。因为 (x_i, y_i) 和 (x_j, y_j) 是同一条直线上的两点,所以它们一定有相同的参数 (p', q') , 而这一点正是参数空间 PQ 中两条直线 $q = -px_i + y_i$ 和 $q = -px_j + y_j$ 的交点。由此可见,图像空间 XY 中过点 (x_i, y_i) 和 (x_j, y_j) 的直线上的每一点,都对应于参数空间中的一条直线,而这些直线必定相交于一点 (p', q') , (p', q') 恰恰就是图像空间 XY 中那条直线方程的参数。这样,通过霍夫变换,可以将图像空间中直线的检测问题转化为参数空间中点的检测问题。而参数空间中点的检测只要进行简单的累加统计就可以完成。

霍夫变换的具体步骤如下:

(1) 在参数空间 PQ 中建立一个二维的累加数组 A 。假设斜率 p 和截距 q 的取值范围分别为 $[p_{\min}, p_{\max}]$ 和 $[q_{\min}, q_{\max}]$, 则累加数组如图 8.18 所示。数组 A 初始化为零。

(2) 对图像空间中的每一个边缘点,让 p 从 p_{\min} 到 p_{\max} 取值,根据式(8.24)得到对应的 q 值,将对应的数组元素 $A(p, q)$ 进行累加。计算结束后,根据 $A(p, q)$ 的值确定在 (p, q) 处共线点的数量。根据 $A(p, q)$ 的最大值所处的位置 (p^*, q^*) , 就可以找到图像空间中的参数为 (p^*, q^*) 的直线。

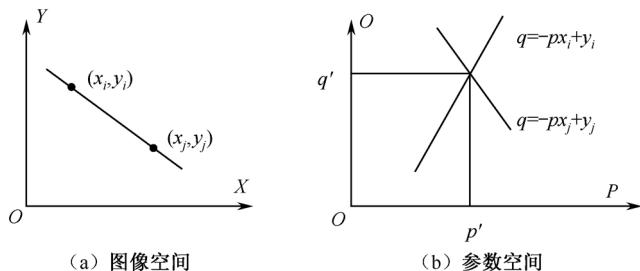


图 8.17 图像空间和参数空间中点和线的对偶性

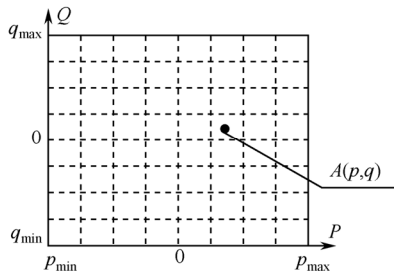


图 8.18 参数空间中的累加数组

综上所述,霍夫变换也可以视为一种聚类分析技术,图像空间中的每一点对参数空间的参数集合进行投票表决,获得多数表决票的参数即为所求的特征参数。上述计算过程中累加数组 A 的大小对计算量和精度要求较高。当被检测直线为垂直时,斜率 p 将为无穷大,计算量激增,采用极坐标可以解决这一问题。

8.5.2 极坐标系中的霍夫变换

设 ρ 为直线到原点的垂直距离, θ 为原点到直线的垂线与 X 轴的夹角, 则极坐标中的点法式直线方程为

$$\rho = x \cos \theta + y \sin \theta \quad (8.25)$$

可以证明, 与直角坐标系中的霍夫变换不同的是, 式 (8.25) 将图像空间 XY 上的点映射为 $\rho\theta$ 平面上的正弦曲线。如图 8.19 所示, 某条直线的参数为 (ρ', θ') , 其上的两点分别为 (x_i, y_i) 和 (x_j, y_j) 。如图 8.20 所示, 变换后两点形成的正弦曲线分别为 $\rho = x_i \cos \theta + y_i \sin \theta$ 和 $\rho = x_j \cos \theta + y_j \sin \theta$, 两共线点生成参数空间中的曲线交于点 (ρ', θ') 。

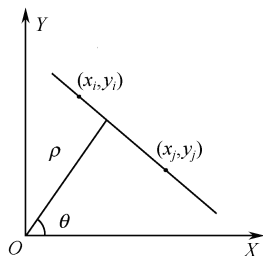


图 8.19 直线的极坐标表示

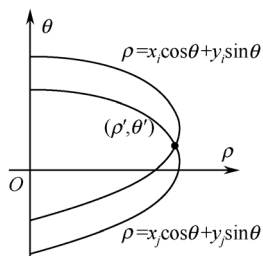


图 8.20 参数空间对应的曲线

在参数空间建立累加二维数组 A 的方法与直角坐标系中的方法类似, 但参数为 ρ 和 θ 。 θ 的取值范围为 $[-90^\circ, +90^\circ]$; 若图像大小为 $M \times N$, 则 ρ 取值范围为 $[-\sqrt{M^2 + N^2} / 2, \sqrt{M^2 + N^2} / 2]$ 。

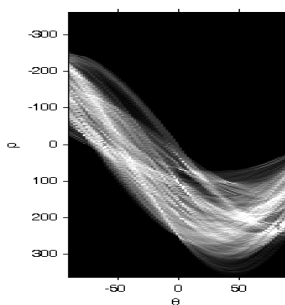
在进行霍夫变换前应该先对图像进行预处理, 一般先对灰度图像进行二值化处理, 然后细化边缘得到图像骨架, 再采用霍夫变换提取图像中的直线。图 8.21 为一对 house 图像进行霍夫变换的实例, 调用 IPT 二值图像标准霍夫变换函数 $[H, \text{THETA}, \text{RHO}] = \text{hough}(\text{BW})$, 其中 BW 是待变换的二值图像, $(\text{THETA}, \text{RHO})$ 分别是离散化的 (θ, ρ) 。H 是霍夫变换矩阵。



(a) house 图像



(b) 二值化图像



(c) 二值化图像的霍夫变换

图 8.21 采用霍夫变换提取图像中的直线

霍夫变换做扩展后可以检测所有给出解析式的曲线, 如圆等。进一步, 利用广义霍夫变换可以检测无解析式的任意形状边界。

8.6 区域生长法

区域生长 (region growing) 是一种串行区域分割算法, 其基本思想是将具有相似性质 (如灰度级、

纹理、颜色等)的像素集合起来构成区域,它们对应于实际感兴趣的目标。值得注意的是,区域生长法计算复杂度较高,较少应用在实时要求高的场合。

区域生长的基本步骤是:先对每个需要分割的区域找一个种子像素作为生长的起点,然后将种子像素周围邻域中与种子像素具有相同或相似性质的像素合并到这一区域中。将这些新像素当做新的种子像素继续进行上面的过程,直到没有满足条件的像素可被包括进来为止。这样,一个区域就长成了。可见,在实际应用区域生长法时需要解决三个问题:

(1) 选择或确定一组能正确代表所需区域的种子像素。通常,这可借助于具体问题的特点进行。例如,在军用红外图像中检测目标时,由于一般情况下目标辐射较大,所以可选用图中最亮的像素作为种子像素。利用迭代的方法从大到小逐步收缩也是一种典型的方法。若对具体问题没有先验知识,则常可借助生长所用准则对每个像素进行相应的计算。如果计算结果呈现聚类的情况,则接近聚类重心的像素可取为种子像素。

(2) 确定在生长过程中能将相邻像素包括进来的准则。生长准则的选取不仅依赖于具体问题本身,也和所用图像数据的种类有关。例如,当图像为彩色时,仅用单色的准则效果就会受到影响。另外还需要考虑像素间的连通性和邻近性,否则有时会出现无意义的分割结果。

(3) 制定让生长过程停止的条件或规则。一般生长过程在进行到再没有满足生长准则需要的像素时停止。但常用的基于灰度、纹理、彩色的准则大都基于图像中的局部性质,并没有充分考虑生长的历史。为增加区域生长的能力,常需考虑一些与尺寸、形状等图像和目标的全局性质有关的准则。在这种情况下,需对分割结果建立一定的模型或辅以一定的先验知识。

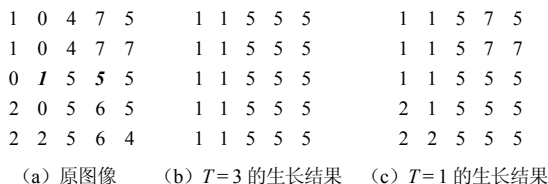


图 8.22 区域生长实例

图 8.22 所示为已知种子点进行区域生长的一个示例。图 8.22 (a) 给出了需分割的图像,设已知有两个种子像素(斜体黑色的两个像素点),先要进行区域生长。假设这里采用的判断准则是:如果所考虑的像素与种子像素的灰度值差的绝对值小于某个门限值,则将该像素包括进种子像素所在的区域。图 8.22 (b) 给出了 $T=3$ 时的区域生长结果,整幅图被较好地分成两个区域;图 8.22 (c) 给出了 $T=1$ 时的区域生长结果,有些像素无法判定。由此例可见门限的选择是很重要的。

由此例可见门限的选择是很重要的。

如图 8.23 所示,图 8.23 (a) 为盆腔骨 CT 原图,图 8.23 (b) 和图 8.23 (c) 分别是采用边缘提取和区域生长方法分割的结果图,可见两种分割方法具有互补作用。

与区域生长法类似的串行区域分割算法还有分裂合并法,基本思想是从整幅图像开始,通过不断地分裂、合并,得到各个不同性质的区域。该算法通常用四叉树或金字塔式分割技术完成。

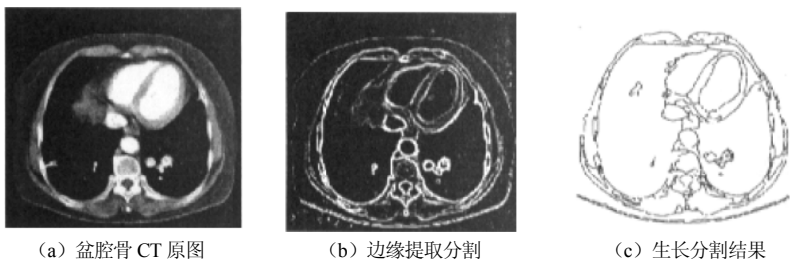


图 8.23 盆腔骨 CT 图像的区域生长分割

烟尘图像分割对于大气污染监控、火灾预警和军事情报获取具有重要意义。以下分别选用了早晨、

中午和傍晚的烟尘图像（见图 8.24）分割结果说明了阈值分割（见图 8.25）、区域生长（见图 8.26）、分裂合并（见图 8.27）三种基于区域的分割方法的特点。可见，阈值分割算法保持了烟尘区域的完整性，但像素的噪声区域较多，边界不够分明。区域生长法分割的边界分明，像素的噪声区域较少，但烟尘区域的完整性欠缺。分裂合并算法兼有上述两算法的优点，但计算的时间较长、空间复杂度较高。

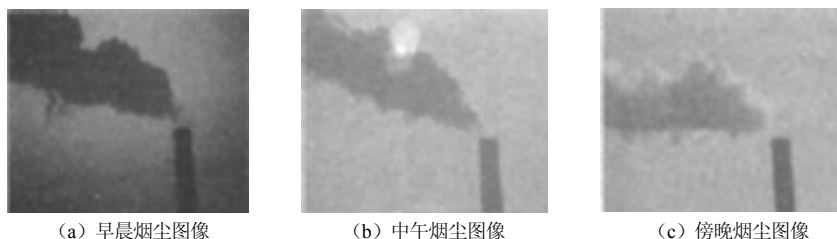


图 8.24 三幅原始烟尘图像

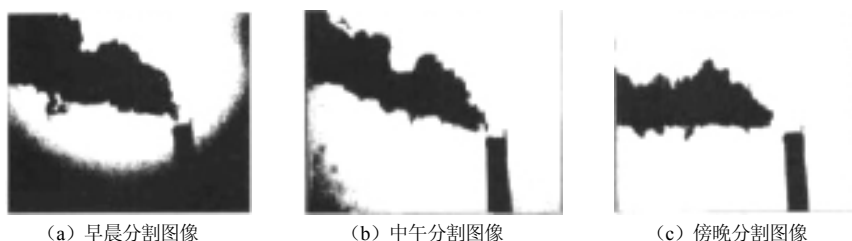


图 8.25 烟尘图像的阈值分割结果

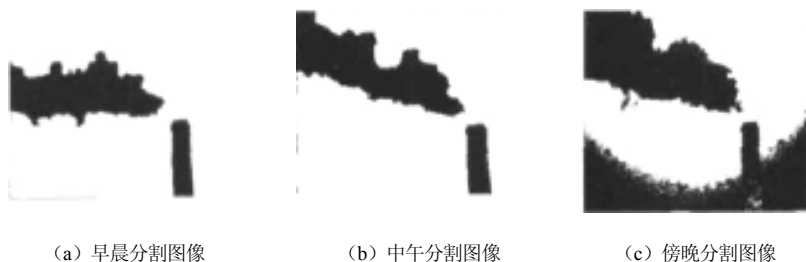


图 8.26 烟尘图像的区域生长结果

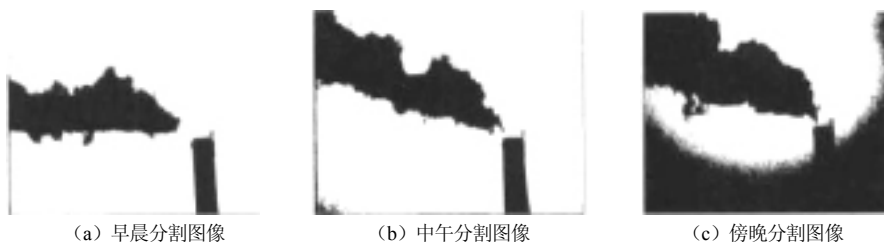


图 8.27 烟尘图像的分裂合并结果

8.7 图像分割方法的比较

8.7.1 边缘检测的优缺点

边缘检测利用不同区域间像素灰度不连续的特点，检测出区域间的边缘，实现图像分割。不同图

像灰度不同，边界处一般会有明显的边缘，利用此特征可以分割图像。通过求微分算子来检测到边缘处像素的灰度不连续值进行边缘检测。边界跟踪是先检测到边缘再串行连接成闭合边界的方法，此方法很容易受起始点的影响。

边缘检测的难点在于边缘检测时抗噪性和检测精度之间的矛盾。若提高检测精度，则噪声产生的伪边缘会导致不合理的轮廓；若提高抗噪性，则会产生轮廓漏检和位置偏差。

8.7.2 区域分割的优缺点

区域分割把具有某种相似性质的像素连通，从而构成最终的分割区域。它利用了图像的局部空间信息，可有效地克服边缘检测存在的图像分割空间不连续的缺点。

基于区域的分割方法往往会造成图像的过度分割，而单纯的基于边缘检测方法有时不能提供较好的区域结构，为此可将基于区域的方法和边缘检测的方法结合起来，发挥各自的优势以获得更好的分割效果。

8.8 实验：图像分割

1. 实验目的

- (1) 熟悉 MATLAB 图像处理工具箱的使用方法。
- (2) 熟悉图像边缘检测的方法。
- (3) 理解边缘检测在图像分割中的应用。
- (4) 了解和掌握 Photoshop 选择区工具。包括矩形选框工具、套索工具、魔棒工具、遮罩工具等。

2. 实验主要仪器设备

- (1) 台式机或笔记本电脑。
- (2) MATLAB（安装了图像处理工具箱）、Photoshop 软件。
- (3) 典型的灰度、彩色图像文件。

3. 实验原理

(1) MATLAB 采用差分、梯度、拉普拉斯算子及各种高通滤波处理方法可以对图像边缘进行检测。检测的方法有梯度算子、拉普拉斯算子、方向算子、坎尼算子，各种方法实现的原理可以参考本章中的相关各节的内容。

(2) Photoshop 选择区工具，包括矩形选框工具、套索工具、魔棒工具、遮罩工具等。参见 2.3.2 节“图像处理软件”。魔棒工具的选择能力依赖于图像边缘的色彩对比度。可以根据不同的颜色特点为魔棒工具设置不同的参数，以达到最好的选择效果。选择方式的使用方法与矩形选框工具一样。容差能够有效地控制魔棒工具的选择灵敏度，用来控制魔棒工具在识别各像素色值差异时的容差范围。取值越大，容差的范围越大；相反，取值越小，容差的范围越小。

4. 实验内容

- (1) 利用 MATLAB 图像处理工具箱提供的 DEMOS 进行实验。
- (2) 利用 MATLAB 图像处理工具箱的边缘检测函数编程进行边缘检测。
- (3) 理解边缘检测在图像分割中的应用，比较各种边缘检测方法。

5. 实验步骤

- (1) 在 MATLAB 的 Command 窗口中，键入 Demo，并执行。

(2) 在 MATLAB 的 Demo 窗口中, 选择 ToolBox 并双击点开, 选择 Image Processing 中的 Morphology, Analysis, and Segmentation Demos。

(3) 运行 Edge Detection, 选择不同图像、不同的检测方法、不同的阈值, 观察检测结果。

(4) 仿照本书例题, 选择不同图像、不同的检测方法、不同的阈值, 编程实现边缘检测并观察检测结果。

6. 实验报告要求

(1) 说明利用 MATLAB 图像处理工具箱进行边缘检测的方法。

(2) 记录某一图像经各种检测方法检测后的图像。

(3) 分析检测结果。

(4) 回答思考题。

(5) 心得和体会。

7. 预习要求

(1) 了解 MATLAB 图像处理工具箱中边缘检测函数的功能。

(2) 了解图像边缘检测的原理。

8. 思考题

(1) 说明图像边缘检测的原理。

(2) 比较不同图像的边缘检测的差别, 并分析原因。

本章小结

图像分割是图像理解和分析的前提和重要组成部分, 图像分割的好坏直接影响图像分析的结果。分割的算法很多, 大致分为边界分割和区域分割, 各种算法的实现程度难易、优缺点不一。边界方法假设图像分割结果的某个子区域在原图像中一定会有边缘存在; 区域方法假设图像分割结果的某个子区域一定会有相同的性质, 而不同区域的像素则没有共同的性质。基于区域的分割方法往往会造成图像的过度分割, 而单纯的基于边缘检测的方法有时不能提供较好的区域结构, 为此可将基于区域的方法和边缘检测的方法结合起来, 发挥各自的优势以获得更好的分割效果。

本章的重点是理解图像分割在图像处理中的地位和作用、各种常见的图像分割方法, 难点是在图像处理时如何选用合适的分割方法以保证图像的特征提取。读者在学习时可以结合实例来理解各种分割方法。

思考题与习题

8.1 举例说明分割在图像处理中的实际应用。

8.2 根据 4 连通或 8 连通准则, 判断图 8.28 所示图像中的目标并编写 MATLAB 程序。

8.3 阈值分割技术适用于什么场景下的图像分割?

8.4 图像的背景亮度均值为 50, 方差为 300; 目标占图像像素总数的百分比为 20%, 均值为 150, 方差为 100。求最佳分割阈值。

8.5 边缘检测的理论依据是什么? 有哪些方法? 各有什么特点?

8.6 设计一个利用 Sobel 算子、Roberts 算子、LoG 算子进行边缘检测的程序, 比较各边缘检测算子检测的视觉效果与运算量。

8.7 说明运用具有微分功能的小波，可通过二进小波变换方便地实现对多个分辨率相似图像的梯度矢量的计算，从而实现多尺度边缘提取。

8.8 为什么要进行边缘跟踪？简述光栅扫描跟踪的基本步骤。

8.9 实际采用霍夫变换检测直线时，为什么不采用 $y = kx + b$ 的表示形式？

8.10 证明：在极坐标系中的霍夫变换将图像空间 XY 的共线点映射为 $\rho\theta$ 平面上的正弦曲线并交汇于一点。

8.11 尝试编写霍夫变换的 MATLAB 程序。

8.12 对于图 8.29 的图像采用基于区域灰度差进行区域增长，给出灰度差值：(1) $T=1$ ；(2) $T=2$ ；(3) $T=3$ 三种情况下的分割图像。

1	1	1	0	0	0	0	0
1	1	1	0	1	1	0	0
1	1	0	0	1	1	0	0
1	1	0	0	0	1	1	0
0	1	0	0	0	0	1	0
1	0	1	0	0	0	1	0
1	1	1	0	0	1	1	0
1	1	1	0	0	0	0	0

图 8.28 待判断目标的图像

1	0	4	7	5
1	0	4	7	7
0	1	5	5	5
2	0	5	6	5
2	2	5	6	4

图 8.29 待区域生长的图像

第9章 形态学图像处理

内容提要

数学形态学基于结构元素的基本思想以及作为基本运算的上、下确界关系表示反映了自然界中一类相当普遍的现象。此方法比其他空域或频域图像处理和分析方法具有一些明显的优势。数字形态学的主要研究内容是图像形态的几何特征,结构特征的定量描述与分析,是线性向非线性处理的延伸。用最基本的移位与逻辑运算的复合操作代替大量复杂的图像处理运算是形态变换最大的特点。本章介绍数学形态学的发展简史、数学形态学运算的基本理论和方法、数学形态学图像处理的基本思想、形态学处理的各种应用。

知识要点

- 数学形态学的发展简史、基本思想和方法。
- 二值形态学的基本理论、方法和算法。
- 灰值形态学的基本理论、方法和算法。
- 彩色形态学的基本理论、方法和算法将结合彩色图像处理技术在下一章介绍。
- 数学形态学图像处理的典型应用(如消除噪声、边缘检测、目标识别等)。

教学建议

- 本章教学安排4学时。若本课程学时较少,本章可以作为选讲或自学内容。
- 重点让读者了解数学形态学的概念、基础理论和基本应用。
- 学习本课程的先修知识包括集合论、矩阵的表示与运算、图像滤波的基本方法等。

9.1 概述

9.1.1 数学形态学的发展简史及基本思想

数学形态学 (Mathematical Morphology) 的历史可回溯到 19 世纪欧拉 (Euler)、20 世纪闵可夫斯基 (Minkowski) 等人的研究。1964 年, 法国的 Matheron 和 Serra 在积分几何的研究成果上, 将数学形态学引入图像处理领域, 并研制了基于数学形态学的图像处理系统。1968 年, 在巴黎矿业学院创建了枫丹白露数学形态学研究中心。Matheron 于 1975 年出版的 *Random Sets and Integral Geometry* 一书中严谨而详尽地论述了随机集合论、积分几何论和拓扑逻辑论, 为数学形态学奠定了坚实的理论基础。1982 年, Serra 出版的专著《图像分析与数学形态学》是数学形态学发展的重要里程碑, 表明数学形态学在理论上趋于完备并且在医学、生物学、遥感、机器人视觉等领域取得了许多重要的成果。1985 年以后, 一些相关领域的国际会议开始把数学形态学列为学术讨论专题, 或专门举行研讨会。国际光学工程学会 (SPIE) 从 1990 年起每年举办一次 *Image Algebra and Morphological Image Processing* 会议。与此同时, 许多数学形态学方面的成果也被大量刊登在有关的国际刊物上, 如 1986 年《计算机视觉与图形图像处理杂志》(GVGIP) 出版了数学形态学专刊, 1989 年和 1994 年国际《信号处理杂志》(*Journal of Signal Processing*) 也出版了形态学在信号处理中的应用研究专辑。数学形态学方面的专著也相继出版, 进一步兴起了形态学的研究热潮。经过半个世纪的发展, 数学形态学无论是在理论方面还是在应用方面都取得了很多成就。

数学形态学是一门建立在严格数学理论基础上的, 分析研究空间结构的形状、框架的学科。它主要以积分几何、集合代数和拓扑论为理论基础, 此外还涉及随机集论、近世代数和图论等一系列数学分支。数学形态学的理论虽然很复杂, 被称为“惊人的数学”, 但它的基本思想却是简单而完美的。数学形态学的基于集合的观点是极其重要的。这意味着:

- (1) 它的运算由集合运算 (如并、交、补等) 来定义。
- (2) 所有的图像都必须以合理的方式转换为集合。

这一基于集合观点的一个自然的结果是: 形态学算子的性能主要以几何方式进行刻画, 传统的理论却以解析的方式来描述, 而几何描述的特点在某些应用场合更适合视觉信息的处理和分析。这一基本思想如图 9.1 所示。

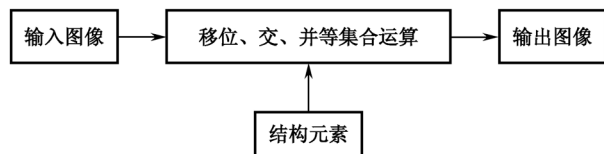


图 9.1 图像处理的数学形态学方法

根据探测研究图像的不同结构特点, 结构元素可携带形态、大小、灰度、色度等信息。不同点的集合形成具有不同性质的结构元素。由于不同的结构元素可以用来检测图像不同侧面的特征, 因此设计符合人的视觉特性的结构元素是分析图像的重要步骤。

最基本的形态学运算有膨胀 (dilation)、腐蚀 (erosion)、开 (opening)、闭 (closing)。用这些算子及其组合来进行图像形状和结构的分析及处理, 可以解决抑制噪声、特征提取、边缘检测、形状识别、纹理分析、图像恢复与重建等方面的问题。

利用数学形态学进行图像处理有其独有的一些特性:

- (1) 它反映的是一幅图像中像素点间的逻辑关系, 而不是简单的数值关系。

形态学图像处理的基本思想是利用结构元素 (structuring element, SE) 作为“探针”在图像中不断移动, 在此过程中收集图像的信息, 分析图像各部分间的相互关系, 从而了解图像的结构特征。结构元素的选择十分重要,

- (2) 它是一种非线性的图像处理方法, 并且具有不可逆性。
- (3) 它可以并行实现。
- (4) 它可以用来描述和定义图像的各种集合参数与特征。

9.1.2 几个基本概念

1. 击中与击不中

设有两幅图像 A 和 B , 如果 $A \cap B \neq \emptyset$ (空集), 那么称 B 击中 (hit) A , 记为 $B \uparrow A$; 否则, 如果 $A \cap B = \emptyset$, 称 B 击不中 (miss) A 。

2. 平移和反射

设 A 是一幅数字图像, a 是 A 的元素 (即 $a \in A$); x 是一个点 (可视为从原点到该点的一个向量 x), 那么定义 A 被 x 平移后的结果为

$$A + x = \{a + x | a \in A\} \quad (9.1)$$

即整个图像沿着向量 x 的方向平行移动。图 9.2 (a) 中的 $\Delta A_1 B_1 C_1$ 是由 ΔABC 平移 x 而成的。

一幅数字图像 A 关于原点的反射定义为

$$A^V = \{a | -a \in A\} \quad (9.2)$$

即反射后的图像 A^V 是由原图像 A 的每个点取相反数后得到的。图 9.2 (b) 中的“像”的反射是经旋转 180° 得到“影”。

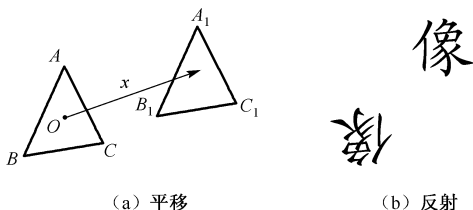


图 9.2 图像的平移与反射

3. 结构元素

结构元素与被处理的目标图像中抽取何种信息密切相关。如前所述, 为了确定目标图像的结构, 必须逐个考察图像各部分之间的关系, 并且进行检验, 最后得到一个各部分之间关系的集合。在考察目标图像各部分之间的关系时, 需要设计一种“结构元素”。在图像中不断移动结构元素, 就可以考察图像之间各部分的关系。

根据不同的图像分析目的, 常用的结构元素有方形、扁平形、圆形等。扁平形结构元素是一种重要的、在其定义域上取常数的结构元素。在多尺度形态学分析中, 结构元素的大小可以变化, 但结构元素的尺寸一般要明显小于目标图像的尺寸。IPT 提供了多个可供选择的创建不同形状和大小结构元素对象的函数 `strel`。根据 `shape`、`parameters` 指定的类型和参数集创建一个结构元素 `SE` 的调用格式

$$SE = \text{strel}(\text{shape}, \text{parameters})$$

创建一个指定邻域的平面结构元素, 其调用格式有

$$SE = \text{strel}(\text{'arbitrary'}, \text{NHOOD})$$

其中, `NHOOD` 为包含 1 和 0 元素的矩阵; 1 元素的位置定义了操作的邻域, `NHOOD` 的原点是其中心元素, 位置在 `floor((SIZE(NHOOD)+1)/2)`。也可以忽略参数串 `'arbitrary'` 而只使用 `strel(NHOOD)`。

【例 8.1】 创建以下的结构元素: ① 斜边长度约为 5, 角度为 45° 的线型结构元素; ② 5×5 大小的正方形结构元素; ③ 半径为 5 的圆盘形结构元素; ④ 四个顶点到中心距离为 3 的钻石型结构元素。

【解】 程序如下:

```
se1 = strel('line',5,45)    %斜边长度约为 5, 角度为 45 度的线型
se2 = strel('square',5)     %5×5 正方形
se3 = strel('disk',4)       %半径为 4 的圆盘形
se4 = strel('diamond',3)    %四个顶点到中心距离为 3 的钻石型
```

运行结果如下：

```
se1 =
Flat STREL object containing 3 neighbors.
Neighborhood:
  0   0   1
  0   1   0
  1   0   0

se2 =
Flat STREL object containing 25 neighbors.
Decomposition: 2 STREL objects containing a total of 10 neighbors
Neighborhood:
  1   1   1   1   1
  1   1   1   1   1
  1   1   1   1   1
  1   1   1   1   1
  1   1   1   1   1

se3 =
Flat STREL object containing 69 neighbors.
Decomposition: 6 STREL objects containing a total of 18 neighbors
Neighborhood:
  0   0   1   1   1   0   0
  0   1   1   1   1   1   0
  1   1   1   1   1   1   1
  1   1   1   1   1   1   1
  1   1   1   1   1   1   1
  0   1   1   1   1   1   0
  0   0   1   1   1   0   0

se4 =
Flat STREL object containing 25 neighbors.
Decomposition: 3 STREL objects containing a total of 13 neighbors
Neighborhood:
  0   0   0   1   0   0   0
  0   0   1   1   1   0   0
  0   1   1   1   1   1   0
  1   1   1   1   1   1   1
  0   1   1   1   1   1   0
  0   0   1   1   1   0   0
  0   0   0   1   0   0   0
```

9.2 二值形态学

二值图像指灰度值只取两种值的图像。为方便起见，这两个灰度值可取为0（相应的点构成背景）和1（相应的点构成景物）。二值形态学处理算法都是以膨胀和腐蚀两种最基本的运算为基础的。膨胀的效果是给图像中的对象边界增加像素，而腐蚀则删除边界的某些像素。在形态学运算中，输出图像

中的像素是由输入图像相应像素及邻域像素按一定规则确定的。下面对几种常用形态学运算的定义及实现方法分别进行介绍。假设集合 A 为图像集合, 集合 B 为结构元素, 数学形态学运算是用 B 对 A 进行操作。

9.2.1 二值腐蚀

集合 A (输入图像) 被集合 B (结构元素) 腐蚀, 表示为 $A \ominus B$, 其定义为

$$A \ominus B = \{x : B + x \subset A\} \quad (9.3)$$

可见, $A \ominus B$ 由将 B 平移 x 仍包含在 A 内的所有点 x 组成。如果将 B 视为模板, 那么 $A \ominus B$ 则由在将模板平移的过程中, 所有可以填入 A 内部的模板的原点组成, 如图 9.3 所示。

IPT 提供了多个可供选择的利用结构元素对象 SE 或 NHOOD 腐蚀图像的函数 `imerode`。创建一个指定邻域的平面结构元素, 其调用格式有

```
J = imerode(I,SE)
J = imerode(I,NHOOD)
J = imerode(I,SE,PACKOUT,M)
J = imerode(...,PADOPT)
```

其中, 图像 I 可以是灰度图像、二值图像或压缩二值图像, J 是返回的已膨胀的图像。PACKOUT 和 PADOPT 是两个优化因子。PACKOUT 取值为

'ispacked' (I 被视为压缩二值图像) 和 'notpacked' (I 被视为正常矩阵, 此为默认值)。PADOPT 取值为 'same' (使输出图像与输入图像大小相同, 此为默认值。如果 PACKOPT 取 'ispacked', 则 PADOPT 必须取 'same') 和 'full' (计算全部腐蚀)。

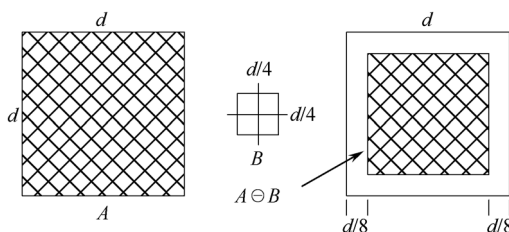


图 9.3 腐蚀示意图

9.2.2 二值膨胀

膨胀是腐蚀运算的对偶运算, 可以通过对补集的腐蚀来定义。我们以 A^C 表示集合 A 的补集, \check{B} 表示 B 关于坐标原点的反射。那么, 集合 A 被集合 B 膨胀表示为 $A \oplus B$, 其定义为

$$A \oplus B = [A^C \ominus \check{B}]^C \quad (9.4)$$

为了利用结构元素 B 膨胀集合 A , 可将 B 相对原点旋转 180° 得到 \check{B} , 再利用 \check{B} 对 A^C 进行腐蚀。膨胀的结果如图 9.4 所示。

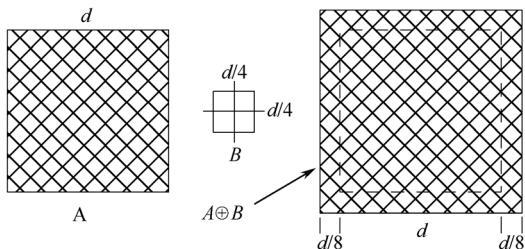


图 9.4 膨胀示意图

MATLAB 图像处理工具箱提供了多个可供选择的利用结构元素对象 SE 或 NHOOD 膨胀图像的函数 `imdilate`。创建一个指定邻域的平面结构元素, 其调用格式有

```
J = imdilate(I,SE)
J = imdilate(I,NHOOD)
J = imdilate(I,SE,PACKOUT)
J = imdilate(...,PADOPT)
```

【例 8.2】 分别 square 和 disk 结构元素膨胀 text 图像。

【解】为了比较两种不同结构元素膨胀的效果，取 2×2 大小的正方形结构元素 $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 和半径为 1

的圆盘形结构元素 $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ ，运行结果显示不同形状和大小的结构元素，其膨胀效果是不一样的。程序如下：

```
I = imread('text.tif');
se1 = strel('square',2);
I1 = imdilate(I,se1);
se2 = strel('disk',1);
I2 = imdilate(I,se2);
subplot(1,3,1),imshow(I)
subplot(1,3,2),imshow(I1)
subplot(1,3,3),imshow(I2)
```

程序运行结果如图 9.5 所示。

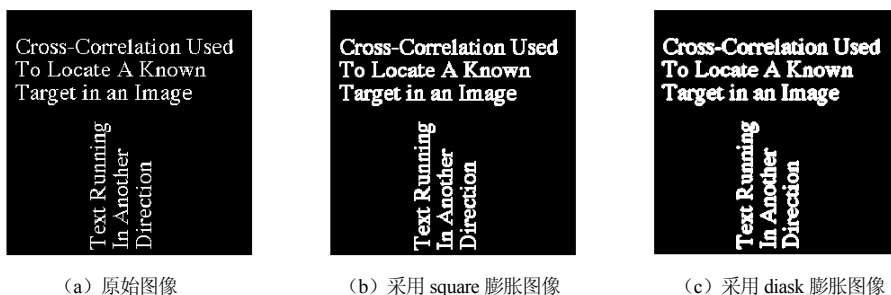


图 9.5 不同结构元素和大小的膨胀效果图

下面以 QR 码为例说明膨胀操作一种应用。QR 码是一种普遍使用的矩阵式二维条码，除具有二维条码所具有的信息容量大、可靠性高、可表示汉字及图象多种信息、保密防伪性强等优点外，还具有以下特点：① 超高速识读，从 QR 是英文“Quick Response（快速反应）”的缩写就不难理解其适宜应用于工业自动化生产线管理等领域；② 全方位识读识读特点；③ 能够有效地表示汉字。QR 码的识别软件甚至可以安装在智能手机上，为用户提供了极大的方便。

通过图像的采集设备，我们得到含有条码的图像，此后主要经过条码定位、分割和解码三个步骤实现条码的识别。条码的定位就是找到条码符号的图像区域，对有明显条码特征的区域进行定位。然后，根据不同条码的定位图形结构特征对不同的条码符号进行下一步的处理。实现条码的定位采用以下步骤：① 对图像进行二值化处理；② 对其二值化图像进行膨胀运算；③ 对膨胀后的图像进行边缘检测得到条码区域的轮廓。找到条码区域后，还要进一步区分矩阵式条码的类型。

QR 码的解码过程如下：得到一幅标准的条码图像后，对该符号进行网格采样，对网格每一个交点上的图像像素取样，并根据阈值确定是深色块还是浅色块。构造一个位图，用二进制的“1”表示深色像素，“0”表示浅色像素，从而得到条码的原始二进制序列值，然后对这些数据进行纠错和译码，最后根据条码的逻辑编码规则把这些原始的数据位流转换成数据码字。

膨胀和腐蚀这两种运算是紧密联系在一起的，一个运算对图像目标的操作相当于另一个运算对图像背景的操作，其对偶性可表示为

$$(A \oplus B)^C = A^C \ominus \overset{\vee}{B} \quad (9.5)$$

$$A \ominus B^C = A^C \oplus \overset{\vee}{B} \quad (9.6)$$

由以上公式和图示可以得出：腐蚀是对图像内部做滤波处理，而膨胀是利用结构元素对图像补集进行填充，因而它是对图像外部做滤波处理。腐蚀具有收缩图像的作用，膨胀具有扩大图像的作用。

图像填充是形态学的一种常用操作，它通过膨胀等操作，根据像素边界求出像素区域。若所用的非边界元素均标为 0，下述过程将把这个区域用 1 来填充：设 B 为对称结构元素，在图像 A 中，从区域边界内的一点 P 开始，将 1 赋给 P ， $X_0 = P$ ，按下列迭代程序填充整个区域。

$$X_k = (X_{k-1} \oplus B) \cap A^C \quad (9.7)$$

式中， $k = 1, 2, 3, \dots$ 结束条件为 $X_k = X_{k-1}$ 。集合 X_k 和 A 的并集包含被填充的集合和它的边界。

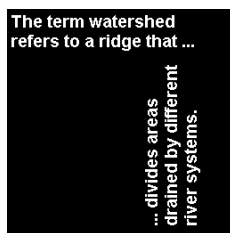
IPT 实现图像区域和孔洞填充操作的函数是 `imfill`。对于二值图像，它可以将相邻背景的像素设置成对象的边界像素；对于灰度图像，它将被较亮区域围绕的黑暗区域的灰度值设置为与围绕区域的像素值相同的数值，在效果上去除了没有连接到边界的部分极小值。实现填充操作有三个步骤：指定填充操作的连通性，指定填充起点，进行填充迭代过程。类似于 Windows 附件“画图”中的“油漆桶”功能，图像 BW1 用泛洪填充算法（flood-fill algorithm）在指定起始位置 LOCATIONS 填充密闭区域的函数调用格式为

$$BW2 = \text{imfill}(BW1, \text{LOCATIONS})$$

【例 8.3】 尝试形态学重构一种操作：用 `imfill` 函数填充 text 图像。

【解】 填充图像 I 的孔洞（holes）形成图像 J。IPT 函数为 $J = \text{imfill}(I, \text{'holes'})$ 。这里孔洞是指较亮像素包围的暗像素区域。程序运行结果如图 9.6 所示。程序如下：

```
I = imread('text.png');
J = imfill(I, 'holes');
subplot(1,2,1), imshow(I);
subplot(1,2,2), imshow(J);
```



(a) 原始图像 text.png



(b) 填充后的图像

图 9.6 填充空洞操作

9.2.3 二值开运算

在形态学图像处理中，除了腐蚀和膨胀这两种基本运算之外，还有两种二次运算起着非常重要的作用，即开运算及其对偶运算闭运算。从结构元素填充的角度看，开运算和闭运算具有更为直观的几何形式。

假设 A 仍为输入图像， B 为结构元素，利用 B 对 A 做开运算，用符号 $A \circ B$ 表示，其定义为

$$A \circ B = (A \ominus B) \oplus B \quad (9.8a)$$

或

$$A \circ B = \bigcup (B + x: B + x \subset A) \quad (9.8b)$$

开运算可以是腐蚀和膨胀的组合运算： A 先被 B 腐蚀，然后再被 B 膨胀。开运算也可以通过计算所有可以填入图像内部的结构元素平移的并集求得。当结构元素 B 扫过整个图像集合内部， $A \circ B$ 就是使结构元素 B 的任何像素不越出图像 A 边界的图像 A 的像素点的集合。

典型的例子是用圆盘 B 对矩形 A 做开运算，结果如图 9.7 所示。开运算具有以下作用：① 利用圆盘做开运算起到磨光内边缘的作用，可以使图像的尖角转化为背景；② 圆盘可以使边界平滑，其圆化作用可以起到低通滤波的效果。③ 断开窄小的连接，消除细小的尖刺保持面积大小不变。

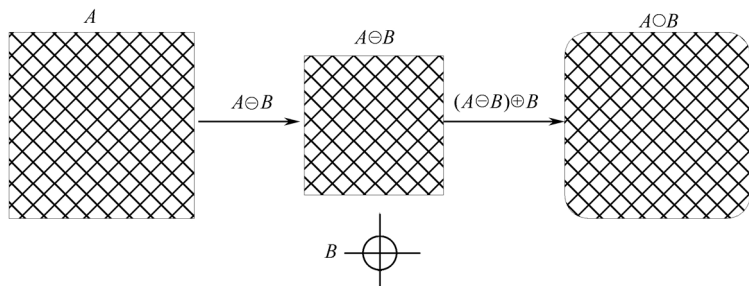


图 9.7 利用圆盘做开运算

IPT 提供了利用结构元素对象 SE 或 NHOOD 对图像 I 进行开运算的函数 `imopen`, 其调用格式为
 $J = \text{imopen}(I, \text{SE})$ 或 $J = \text{imopen}(I, \text{NHOOD})$

9.2.4 二值闭运算

闭运算是开运算的对偶运算, 定义为先做膨胀后做腐蚀。对图像 A 用结构元素 B 做闭运算用符号

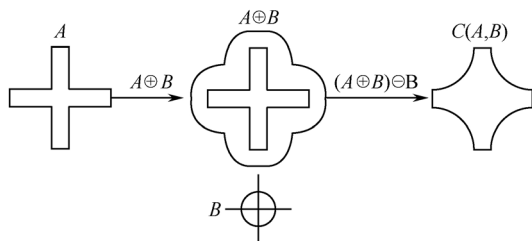


图 9.8 利用圆盘做闭运算

$A \cdot B$ 表示, 其定义为

$$A \cdot B = (A \oplus B) \ominus B \quad (9.9a)$$

或

$$A \cdot B = (A^c \oplus B)^c \quad (9.9b)$$

从上式可以看出: 对图像 A 用结构元素 B 做闭运算可得到一个集合, 该集合中包含所有这样的点 x , x 被一个平移的镜像结构元素覆盖的同时, 平移的镜像结构元素与 A 图像必有一些公共点。由此看出, 初始图像 A 是包含在闭运算后的 $A \cdot B$ 中的, 即闭运算是具有延伸性的运算。图 9.8 描述了闭运算的过程及结果。显然, 闭运算对图像的外部做滤波处理, 仅仅磨光了图像内部的尖角。可见, 开运算具有磨光图像外边界的作用, 闭运算具有磨光图像内边界的作用。IPT 提供了利用结构元素对象 SE 或 NHOOD 对图像 I 进行闭运算的函数 `imclose`, 其调用格式为

$$J = \text{imclose}(I, \text{SE}) \quad \text{或} \quad J = \text{imclose}(I, \text{NHOOD})$$

9.3 灰值形态学

在灰度图像形态处理中, 输入和输出的图像都是灰度级形式的, 这意味着输入和输出像素值在最低灰度值到最高灰度值之间。

9.3.1 灰值腐蚀

形态学源于填充的概念, 而灰值形态学处理的对象是图像信号波形的拓扑特性, 自然结构元素也是一个信号。在二值形态学中, 集合的交、并运算起到关键作用, 而在灰值形态学中这两种运算对应于极大和极小运算。因此, 我们可利用填充、极大/极小概念直接定义灰值运算。

用结构元素 g 对输入图像 $f(x, y)$ 进行灰值腐蚀记为 $f \ominus g$, 其定义为

$$(f \ominus g)(s, t) = \min \{f(s+x, t+y) - g(x, y) \mid s+x, t+y \in D_f, x+y \in D_g\} \quad (9.10)$$

式中, D_f 和 D_g 分别是 f 和 g 的定义域。

为简要说明式 (9.9) 的含义, 用一维函数对其进行简化, 定义为

$$(f \ominus g)(s) = \min \{f(s+x) - g(x) \mid s+x \in D_f, x+y \in D_g\} \quad (9.11)$$

由于结构元素必须在信号的下方, 故空间平移结构元素的定义域必为信号定义域的子集, 否则腐蚀在该点没有定义。实质上结构元素从信号的下面对信号产生滤波作用, 这与结构元素从内部对二值图像滤波的情况是相似的。

灰值腐蚀可按下面的步骤完成: 对于结构元素 g 的定义域 D_g 中的每一点 x , 将信号 f 平移 $-x$, 然后再从每次平移信号之中减去 $g(x)$, 对于结构元素定义域中的每一点都得到一个信号, 对所有这些信号逐点取其最小值, 便可以得到腐蚀结果。图 9.9 描述了灰值腐蚀运算的过程及结果。

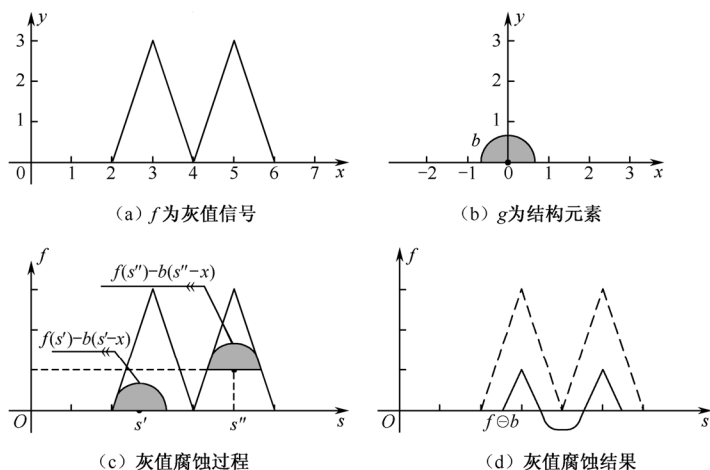


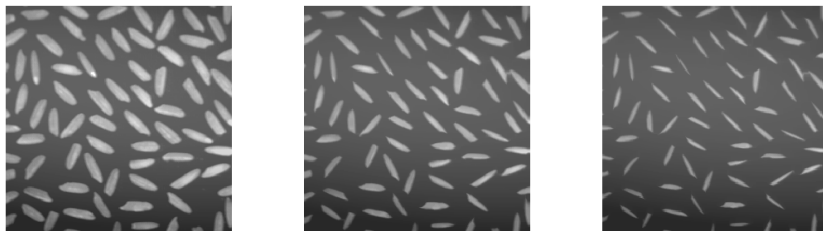
图 9.9 灰值腐蚀运算

【例 8.4】 分别用不同大小的 square 结构元素腐蚀 rice 图像。

【解】 为了比较两种不同大小元素腐蚀的效果, 取 4×4 、 6×6 大小的正方形结构元素, 运行结果显示用不同大小的结构元素膨胀图像的效果是不同的。程序如下:

```
I = imread('rice.tif');
se1 = strel('square',4)
I1 = imerode(I,se1);
se2 = strel('square',6)
I2 = imerode(I,se2);
subplot(1,3,1), imshow(I)
subplot(1,3,2), imshow(I1)
subplot(1,3,3), imshow(I2)
```

程序运行结果如图 9.10 所示。



(a) 原始图像 (b) 采用较小的 square 腐蚀图像 (c) 采用较大的 squarediask 腐蚀图像

图 9.10 不同大小结构元素的腐蚀效果图

9.3.2 灰值膨胀

灰值膨胀可用灰值腐蚀的对偶运算来定义。在定义灰值腐蚀时,我们可以采用求极大值的方法,即在位于信号下方的条件下,求上推结构所能达到的最大值。这里我们利用结构元素的反射,求将信号限制在结构元素的定义域内时,上推结构元素使其超过信号时的最小值来定义灰值膨胀。

用结构元素 g 对输入图像 $f(x, y)$ 进行灰值膨胀记为 $f \oplus g$, f 被 g 膨胀可逐点地定义为

$$(f \oplus g)(s, t) = \max \{f(s-x, t-y) + g(x, y) | s-x, t-y \in D_f, x+y \in D_g\} \quad (9.12)$$

式中, D_f 和 D_g 分别是 f 和 g 的定义域。

为简要说明式 (9.11) 的含义,用一维函数对其进行简化,定义为

$$(f \oplus g)(s) = \max \{f(s-x) + g(x) | s-x \in D_f, x \in D_g\} \quad (9.13)$$

灰值膨胀可按下面的步骤完成:对结构元素 g 的定义域 D_g 中的每一个点 x 将信号 f 平移 x , 然后再对每次平移信号的值加上 $g(x)$, 这样对于结构元素定义域中的每个点都得到一个信号, 对所有这些信号逐点取其最大值, 便可得到膨胀结果。图 9.11 描述了灰值膨胀运算的过程及结果。灰值信号 f 和结构元素 g 与图 9.8 中的相同。

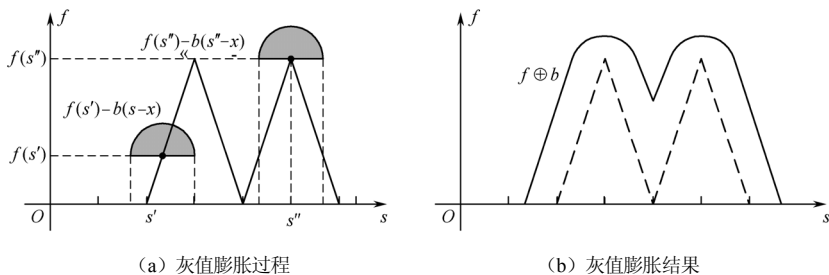


图 9.11 灰值膨胀运算

9.3.3 灰值开运算

具备了腐蚀和膨胀这两种基本灰值形态学运算, 我们便可以定义它们的组合运算——灰值开和灰值闭运算。与二值情况相同, 这两种运算也为对偶运算, 并且两者都可用填充概念来说明。灰值开运算可以参照二值情况来定义, 即先做腐蚀再做膨胀的迭代运算:

$$f \circ g = (f \ominus g) \oplus g \quad (9.14)$$

图 9.12 描述了灰值开运算的过程及结果。

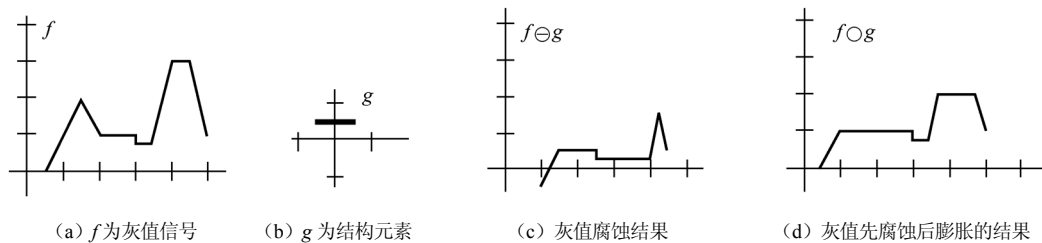


图 9.12 灰值开运算

9.3.4 灰值闭运算

根据对偶性定义, 灰值闭运算定义为

$$f \cdot g = (f \oplus g) \ominus g \quad (9.15)$$

图 9.13 描述了灰值闭运算的过程及结果。

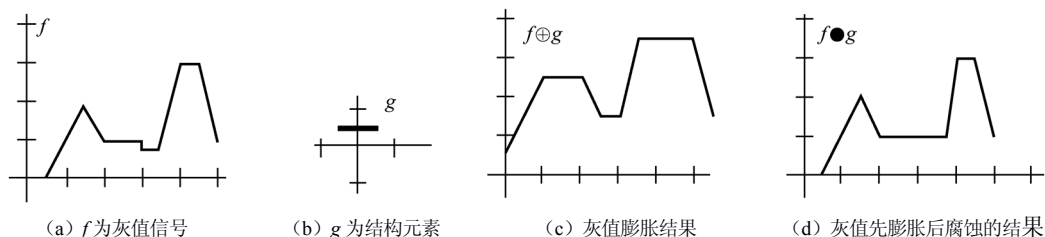


图 9.13 灰值闭运算

从图 9.12 可以看出, 灰值开运算具有非扩展性, 即滤波结果总位于原始图像的下方。它从图像的下方磨光图像灰值表面上突出的尖峰 (即波峰)。从图 9.13 可以看出, 灰值闭运算具有扩展性, 即滤波结果总位于原始图像的上方。它从图像的上方磨光图像灰值表面上向下突出的尖峰 (即波谷)。

图 9.14 所示为细胞组织图像的灰值形态运算: 灰值腐蚀、膨胀、开、闭操作。

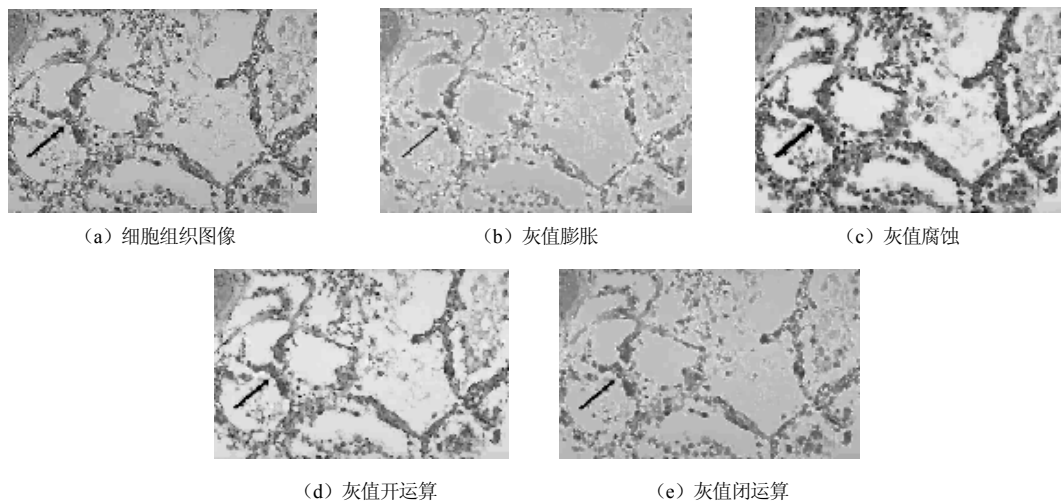


图 9.14 细胞组织图像的灰值形态运算

9.3.5 灰值形态学梯度

利用梯度进行边缘检测的基本原理是: 如果在图像某点出现的梯度值大, 则表示在该点处图像的明暗变化剧烈, 从而可能有边缘存在。在灰值形态学中, 利用扁平结构元素 g 对 f 做腐蚀和膨胀可以得到 f 的局部极大和极小值, 故与数字差分定义的梯度相对应, 一种形态学梯度的定义为

$$\text{GRAD}(f) = (f \oplus g) - (f \ominus g) \quad (9.16)$$

式中, g 为以原点为中心的扁平结构元素。

为了更好地获得边缘检测的效果, 与差分梯度的情况类似, 可以将形态学梯度与阈值结合起来使用。图 9.15 为形态学梯度算法检测到的脑部核磁共振图片的边缘图像。基于数学形态学的边缘信息提取处理优于基于微分运算的边缘提取算法, 它不像微分算法对噪声那样敏感, 提取的边缘也比较平滑。

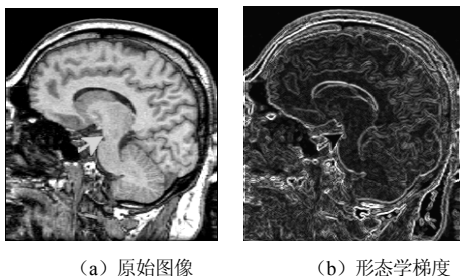


图 9.15 核磁共振图像的形态学梯度

9.3.6 高帽变换和低帽变换

通过这两种变换可以得到灰度图像中一些重要的标记点。例如,在较亮的背景中求暗的像素点或在较暗的背景中求亮的像素点,检测受到噪声污染图像中的边缘等。为了使上述效果更明显,对变换后的图像也可以做阈值处理。

高帽变换 (top-hat transform) 指从一幅原始图像 f 中减去对其做开运算后得到的图像 $THT(f)$, 其定义为

$$THT(f) = f - (f \circ g) \quad (9.17)$$

式中, g 为结构元素。

高帽变换是一种波峰检测器,它在较暗的背景中求亮的像素点很有效。与它相对偶的算子是低帽变换 (bottom-hat transform), 定义为

$$BHT(f) = (f \bullet g) - f \quad (9.18)$$

显然,低帽变换是一种波谷检测器,适合于在较亮的背景中求暗的像素点。

图 9.16 说明了采用扁平结构元素 g , 我们可以利用高帽变换求出图像 f 中的峰值,而利用低帽变换求出图像 f 中的谷值。

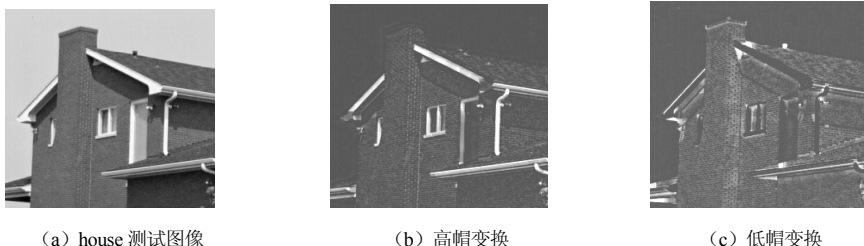


图 9.16 利用高帽和低帽变换检测图像峰值和谷值

9.3.7 开-闭运算和闭-开运算

形态开启和闭合运算被作为最基本的形态滤波运算,但在实际的图像处理中,仅仅采用形态开和闭的滤波效果往往不能令人满意。此时就需要在基本的形态开、闭运算的基础上设计出形态开-闭和形态闭-开组合滤波器,发挥其更好的滤波性能。

形态开-闭运算能够滤除图像中形状小于结构元素的亮噪声,形态闭-开运算可以滤除图像中形状小于结构元素的暗噪声。它们互为对偶操作,定义如下。

开-闭运算:

$$f \square g = f \circ g \bullet g \quad (9.19)$$

闭-开运算:

$$f \blacksquare g = f \bullet g \circ g \quad (9.20)$$

在图像复原处理中,基于数学形态学的形状滤波器可借助于先验的几何特征信息,利用形态学算子有效地滤除噪声,又可以保留图像中的原有信息。

9.4 实验: 形态学图像处理

1. 实验目的

- (1) 了解数学形态学的基本思想和方法。

- (2) 掌握形态学基本运算。
- (3) 学习设计形态学图像处理算法。

2. 实验主要仪器设备

- (1) 台式机或笔记本电脑。
- (2) MATLAB 软件 (含图像处理工具箱)。
- (3) 典型的灰度图像文件。

3. 实验原理

- (1) 最基本的形态学运算有膨胀、腐蚀、开、闭。

(2) 去噪算法: 加入椒盐噪声, 形式如 $J = \text{imnoise}(I, 'salt \& pepper', 0.04)$, 将开和闭结合起来 (开-闭运算或闭-开运算) 构成形态学噪声滤波器, 尝试用不同半径 r 的圆盘形 $se = \text{strel}('disk', r)$ 观察去噪效果。

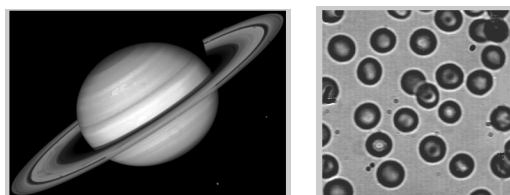
(3) 二值图像的 4 连接 (结构元素为 4 个顶点到中心距离为 1 的钻石型: $se = \text{strel}('diamond', 1)$) 和 8 连接 (结构元素为 3×3 正方形: $se = \text{strel}('square', 3)$) 边缘提取。

4. 实验内容

- (1) 用数学形态学运算去除图像噪声。
- (2) 用形态学梯度 (已膨胀图像减已腐蚀图像) 来提取二值图像中对象的边界像素。

5. 实验步骤

- (1) 选择图 9.17 所示的原始图像作为测试图像。
- (2) 根据设计要求选择合适的去噪算法。
- (3) 显示原始图像和去噪后的图像。
- (4) 根据设计要求提取 4 连接或 8 连接图像边界。
- (5) 显示原始图像和边界图像。



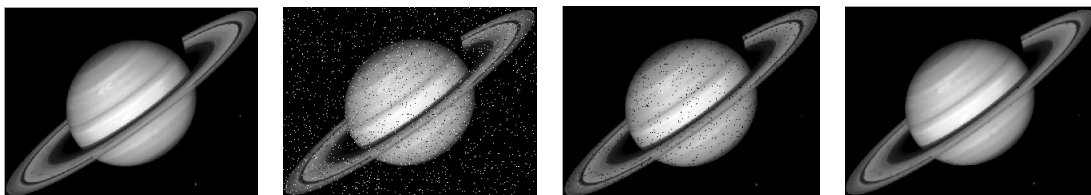
(a) saturn.tif

(b) blood1.tif

图 9.17 供测试的图像

6. 实验报告要求

- (1) 说明形态学算法。
- (2) 原始图像和去噪后的图像, 如图 9.18 所示。



(a) 原始图像

(b) 加入噪声后的图像

(c) 开运算后的图像

(d) 闭运算后的图像

图 9.18 原始图像和去噪后的图像

- (3) 原始图像和边界图像, 如图 9.19 所示。
- (4) 回答思考题。
- (5) 心得和体会。

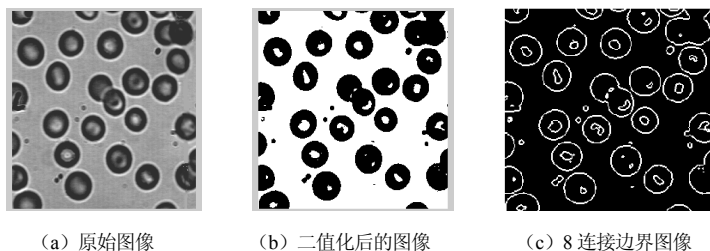


图 9.19 原始图像和边界图像

7. 预习要求

- (1) 了解数学形态学的基本原理和运算。
- (2) 掌握 MATLAB 图像处理工具箱中有关数学形态学图像处理的函数。
- (3) 画出所采用算法的流程框图, 编写相应的程序。

8. 思考题

- (1) 数学形态学基本运算之间有何联系? 各有什么特点?
- (2) 数学形态学用于去噪时, 结构元素形状和大小与噪声点之间有何关系?
- (3) 说明 4 连接和 8 连接检测边界的差异。

本章小结

数学形态学是图像处理和图像分析的有力工具, 数学形态学算子的性能主要以更适合视觉信息处理和几何方式进行刻画, 而传统的理论却以解析方式描述算子的性能。例如, 利用数学形态学方法提取的图像骨架比较连续, 断点少。本章主要介绍了以下主要内容:

- (1) 数学形态学的发展简史、研究方法、特点和应用。
- (2) 描述了二值形态学、灰度形态学的基本运算过程和结果。

学习本章要求掌握二值形态学的基本原理和方法, 了解灰度形态学的概念, 认清两种形态学之间的相互关系。

本章重点要求明确二值数学形态学的概念和应用。

思考题与习题

9.1 数学形态学主要包括哪些研究内容?

9.2 基于数学形态学的图像处理有何特点?

9.3 研究膨胀操作与卷积运算的相似性。定义膨胀 $A \oplus B = \{x: [(\check{B})_x \cap A] \subseteq A\}$, 其中 $(\check{B})_x$ 表示 B 反射后平移 x 。

9.4 画出用一个半径为 $r/4$ 的圆形结构元素腐蚀一个 $r \times r$ 的正方形的示意图。

9.5 编写一个完整的程序, 实现灰值图像的腐蚀、膨胀、开和闭运算, 并对一幅图像进行处理, 研究其结果。

9.6 设计一个算子可以同时检测图像中的峰和谷。

9.7 讨论数学形态学图像处理的主要应用。进一步查找资料, 写一篇关于你感兴趣的应用方面的短文。

第 10 章 彩色图像处理

内容提要

本章主要介绍彩色图像处理的一些基本知识。较为详细地说明了彩色图像的颜色空间变换的概念及相关公式，并结合具体的示例图片，对颜色空间量化和抖动技术进行了概述，介绍了假彩色处理和彩色图像增强的概念和应用，最后探讨了彩色图像形态学的基本原理，并用于彩色图像滤波。彩色图像压缩也是彩色图像存储和传输研究的重要内容，考虑 5.6 节“图像压缩编码标准”中对 JPEG 和 JPEG2000 已做较详细的介绍，请读者结合本章的知识进一步深化。

知识要点

- 颜色空间的表示（含 RGB、Munsell、HSV、HSI、YUV、YIQ、CMYK、 $L^*a^*b^*$ 等模型）。
- 颜色空间的转换。
- 颜色空间的量化。
- 抖动技术。
- 假彩色处理。
- 真彩色增强。
- 伪彩色增强。
- 彩色图像形态学。
- 彩色图像压缩（参考 5.6 节“图像压缩编码标准”）。

教学建议

- 建议学时数为 4 学时。
- 本章的先修知识主要有光学、人类颜色视觉和色度学、线性代数、图像的量化、数学形态学等。
- 要求了解彩色图像处理的基本概念和一些常用的彩色图像处理技术。
- 学会用 MATLAB 工具对彩色图像进行简单处理，如进行颜色空间的变换、彩色图像增强等。
- 结合第 5 章深化彩色图像压缩知识，形成较完整的彩色图像处理知识体系。

10.1 彩色图像处理的基本问题

前面几章以灰度图像处理为主，也涉及部分彩色图像处理的概念，如彩色视觉、彩色图像采集、JPEG 彩色图像压缩原理与标准，这些为今后学习彩色图像处理奠定了必要的知识基础。在许多场合下，灰度图像是彩色图像的一种简化表示，反映了彩色图像的明亮度的特征。客观世界是多彩的，而彩色图像必须用多维变量来表达，颜色信息应该得到更大的关注。从这两方面看，彩色图像处理理论和技术充分利用视觉生理学、视觉心理学和视觉认知学等学科的成果，特别在电视广播、机器人视觉、生物测定学、电子医疗设备、遥感测绘中的多光谱图像处理、基于颜色的网络图像内容检索、智能交通、公安视频监控及安全检查和军事侦察等方面具有重要的理论研究和实际应用意义。

从处理方法分，彩色图像处理可以分成三类：① 颜色空间的转换；② 单独彩色平面的空间处理；③ 彩色向量处理。在某些情况下，单独彩色平面的空间处理与彩色向量处理，都会得出相同的处理结果，如求彩色图像的邻域平均运算。本章主要讨论对各个彩色平面进行空间邻域滤波。

而从应用背景分，彩色图像处理可以在更宽的范围分成多种类别，如彩色图像增强、彩色图像压缩和编码、彩色图像分割与边缘检测、彩色图像形态学等。几乎灰度图像处理需要研究的问题，都可以扩展到彩色图像领域。

作为彩色图像处理的特有问题，如彩色空间和距离、颜色恒常性（color constancy）、彩色静态和动态场景分析、彩色目标跟踪、基于颜色的网络图像内容检索、彩色图像融合和伪彩色技术等，都是值得重视的研究内容。

10.2 颜色空间的表示及其转换

实际应用中常用的颜色空间（color space）很多，有 RGB、HSV、HSI、YUV、YIQ、CMYK 等。目前常用的颜色空间可分为两类，一类是面向硬设备的，如彩色显示器、打印机等，另一类面向以彩色处理为目的的应用。面向硬设备的最常用的颜色空间是 RGB 颜色空间、印刷四色模式（即 CMYK 模型），而面向颜色处理的最常用颜色空间是 HSI 颜色空间及 HSV 颜色空间。下面介绍几种常用的颜色模型（color model），这些模型规定了颜色的建立、描述和观察方式。颜色模型都是建立在三维空间中的，所以与颜色空间密不可分。

各种色彩空间只是同一物理量的不同表示法，因而它们之间存在着相互转换关系。

10.2.1 RGB 模型

RGB 模型采用三原色光模式，用三维空间中的一个点来表示一种颜色，如图 10.1 所示，每个点有 3 个分量，分别代表该点颜色的红、绿、蓝亮度值，亮度值限定在 $[0, 1]$ 。

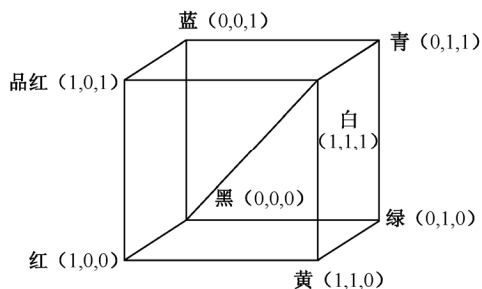


图 10.1 RGB 模型坐标

在 RGB 彩色模型的原点上，任一基色均没有亮度，即原点为黑色。三基色都达到最高亮度时则表现为白色。亮度较低且等量的 3 种基色产生灰色。所有这些点均落在彩色立方体的对角线上，该对角线被称为灰色线。彩色立方体中有 3 个角对应于三基色——红色、绿色和蓝色。余下的 3 个角对应于三补色（three complementary colors）——黄色、青色（蓝绿色）和品红色（紫色）。

图 10.2 按照 RGB 模型将彩色图像分解成了 3 个基色图像分量。

图 10.2 图像的 R 、 G 、 B 分解

RGB 色彩空间有 AdobeRGB、AppleRGB、sRGB 等几种模式，这些 RGB 色彩空间大多与图像外设相关联。Adobe RGB 与 sRGB 则最为常见，也是目前数码相机中重要的设置。sRGB (standard Red Green Blue) 提供了一种标准方法来定义使显示、打印和扫描等各种外设与应用软件彩色语言的协议。对于高版本的 Windows 和 Windows NT，它是默认的色彩模式，它也是 HTML 3.2 的标准色彩模式。

10.2.2 Munsell 模型

在孟塞尔颜色系统 (Munsell Color System) 中，某种颜色可以唯一地用一个色调 (H)、色纯度 (C , Chroma) 及亮暗值 (V , Value) 的颜色片来表示，如图 10.3 所示。色调沿圆周分成 10 个区域，其中 5 个是主色调，另外 5 个是中间色调，分别是红 (R)、红黄 (YR)、黄 (Y)、黄绿 (GY)、绿 (G)、蓝绿 (BG)、蓝 (B)、蓝紫 (PB)、紫 (P) 和紫红 (RP)。相邻的两个位置之间再均分为 10 份，共 100 份。色纯度表示了色的浓淡，从中心 (数值为 0) 向外逐渐增强。普通的颜色实际上限为 10 左右，而反光、荧光等材料可高达 30。颜色的亮暗分成 11 个等级，记为 0~10 级，其中 0 级对应黑，而 10 级对应白。

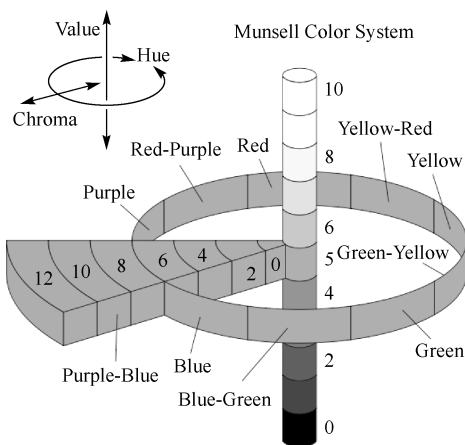


图 10.3 Munsell 彩色空间

Munsell 颜色空间具有以下特点：

- (1) 坐标之间的心理感知独立性，即可以独立感知各颜色分量的变化。
- (2) 线性伸缩性。可感知的颜色差与颜色分量的相应样值上的欧几里得测度之间的距离成比例。
- (3) Munsell 空间在感知上并不是均匀的，也不能直接根据加色原理进行组合。

10.2.3 HSV 模型

在许多实用系统中，大量应用的是 HSV 模型，这个模型是由色调 (H , Hue)、饱和度 (S , Saturation)、亮度 (V , Value) 3 个分量组成的，与人的视觉特性比较接近，具有较强的感知度。RGB 加法原色和 CMYK 减法原色模型，以原色组合的方式定义颜色，而 HSV 以人类更熟悉的方式反映了色相、色纯度和明度三个颜色信息。该模型的重要性在于：一方面消除了亮度成分 V 在图像中与颜色信息的联系，另一方面色调 H 和饱和度 S 分量与人的视觉感受密切相关。基于人的视觉系统的颜色感觉特性，这些特征使 HSV 模型成为一个研究图像处理的重要工具。

HSV 颜色模型用 Munsell 三维空间坐标系统表示。色调 (H) 用角度来标定，在径向方向上饱和度和 (S) 的深浅用离开中心线的距离表示，亮度 (V) 用垂直轴表示。色调 H 用角度 $0^\circ \sim 360^\circ$ 度量，饱和度 S 用百分比来度量，从 0% 到完全饱和的 100%。由于色调 H 是以圆周上的角度度量的，因此由 H 和 S 分量可以构造一个颜色轮 (color wheel)。在颜色轮上，主要颜色沿一个圆均匀分布，

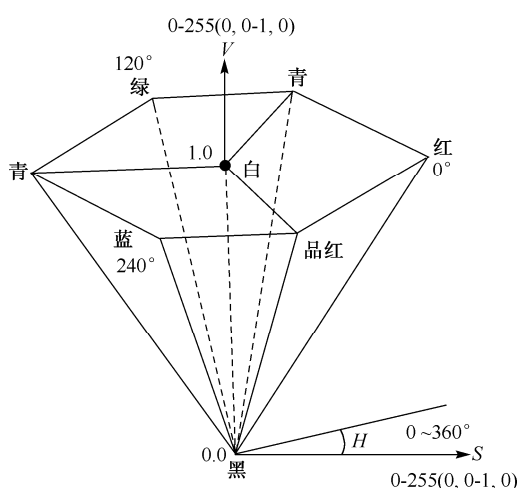


图 10.4 HSV 颜色模型

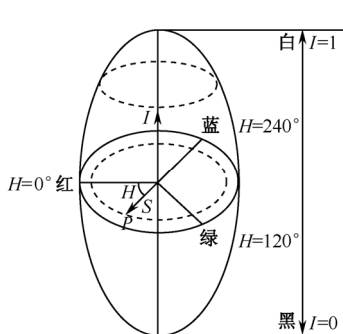
次要颜色位于主要颜色之间。例如绿色位于黄色和青色之间，因此用黄色和青色可以产生绿色。每种颜色与它的补色直接在轮上相对。亮度 V 也通常用百分比度量，从 0%（黑）到 100%（白）。HVS 的坐标系统是圆柱坐标系统，但一般用六棱锥表示，如图 10.4 所示。

HSV 颜色空间目前广泛应用于计算机图形学、科学计算可视化等领域。它是一个均匀颜色空间。在所有的颜色空间中，HSV 模型对应于画家配色模型，比后面将要讨论的 HSI 模型更接近于人对颜色的主观感受，因此很适合基于颜色的图像相似度比较。

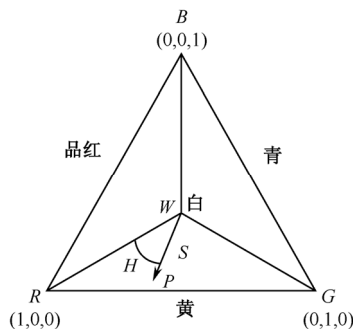
10.2.4 HSI 模型

HSI 色彩空间从人的视觉系统出发，将颜色表示成色调 (H)、饱和度 (S) 和强度 (I)，这 3 个分量构成了彩色空间。色调 (H) 和饱和度 (S) 的含义与 HSV 系统一致，而强度 (I , intensity) 对应于颜色的亮度或灰度。由于人的视觉对亮度的敏感程度远强于对颜色浓淡的敏感程度，为了便于色彩处理和识别，人的视觉系统经常采用 HSI 色彩空间，它比 RGB 色彩空间更符合人的视觉特性。在图像处理和计算机视觉中的大量算法都可在 HSI 色彩空间中方便地使用，它们可以分开处理而且是相互独立的。因此，在 HSI 色彩空间可以大大简化图像分析和处理的工作量。

HSI 彩色模型如图 10.5 (a) 所示，HSI 色彩空间可以用一个圆锥空间模型来描述。用这种描述 HSI 色彩空间的圆锥模型相当复杂，但确能把色调、强度和饱和度的变化情形表现得很清楚。而图 10.5 (b) 显示的是标准 HSI 三角形，三角形的顶点代表了 3 个归一化的彩色分量 (R 、 G 、 B) 的三角系数。色调 H 定义为颜色点 P 至中心的线段与 R 轴之间的夹角。



(a) HSI 彩色模型坐标系统



(b) HSI 彩色三角形

图 10.5 HSI 彩色模型

10.2.5 YUV 模型

在计算机图像处理中，YUV 颜色模型在使用的广泛性方面仅次于 RGB 模型。在彩色电视系统中，采用的就是 YUV 色彩空间。

由于人眼对于亮度的敏感程度大于对于色度的敏感程度，所以完全可以让相邻的像素使用同一个

色度值, 而人眼的感觉不会引起太大的变化。YUV 的基本思想是通过压缩色度信息来达到节省存储空间的目的。我们可以定义出许多 YUV 的格式, 例如相邻两个像素使用一个色度值的 YUV, 而 JPEG 和 MPEG 中相邻 4 个像素使用一个色度值的 YUV12 等。

同一颜色可以用不同的彩色空间表示, 自然可以相互转换。MATLAB 提供了相应的转换函数, 为用户将一种彩色空间的颜色转换为其他空间的颜色表示提供了极大的方便。

除了上述的彩色模型, 还有一些其他的彩色模型, 如 NTSC、YCbCr 模型等。

NTSC 模型广泛应用于美国等国家的电视信号。它的特点是信号的强度信息和颜色信息相分离, 同一个信号可以方便地同时表示彩色图像和黑白图像。在 NTSC 格式中, 图像由 3 个分量表示: 亮度用 Y 表示; 色度用 I 表示; 饱和度用 Q 表示。第 1 个分量是亮度 Y , 表示灰度信息, 后两个分量表示彩色信息。本章的实验指导中给出了将 RGB 空间转换成 YIQ 空间的公式, 即式 (10.17)。MATLAB 提供了 RGB 空间和 YIQ 空间相互转换的函数: RGB2NTSC 函数和 NTSC2RGB 函数。

YCbCr 模型广泛应用于数字视频。在 YCbCr 模型中, Y 为亮度, Cb 和 Cr 共同描述图像的色调, 其中 Cb、Cr 分别为蓝色分量和红色分量相对于参考值的坐标。本章的实验指导中给出了将 RGB 空间转换成 YcbCr 空间的公式 (10.18)。MATLAB 提供了 RGB 空间和 YCbCr 空间相互转换的函数: RGB2YCBCR 函数和 YCBCR2RGB 函数。

10.2.6 CMYK模型

与 RGB 模型不同, CMYK模型是一种适合印刷等用途的减色色彩模式。CMYK 代表印刷上用的 4 种颜色, C 代表青色 (Cyan), M 代表品红色 (Magenta), Y 代表黄色 (Yellow), K 代表黑色 (Black)。如图 10.6 所示, CMY 相交的混色应该为黑色。在实际应用中, 由于青色、品红色和黄色很难叠加形成真正的黑色, 因此引入了黑色 K, 加深暗部色彩。

由于 RGB 和 CMY 互为补色, 因此, 简单的转换关系为

$$R = 1 - C \quad (10.1a)$$

$$G = 1 - M \quad (10.1b)$$

$$B = 1 - Y \quad (10.1c)$$

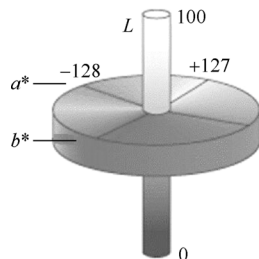


图 10.6 $L^*a^*b^*$ 模型

10.2.7 $L^*a^*b^*$ 模型

$L^*a^*b^*$ 模型基于人眼对颜色的线性 (或均匀) 感知的要求, 感知上的线性意味着在色彩空间上相同数量的变化应当产生大致相同视觉重要性的变化。CIE 1976 (L^*, a^*, b^*) 色彩空间也称为 CIELAB, 它的坐标实际上是 L^*, a^* 和 b^* 。 $L^*a^*b^*$ 空间中的数值描述了人眼能够看到的所有颜色, 且处理速度与 RGB 模型同样快。它是 CIE 发布的一个既不依赖光线 (如 RGB 模型), 也不依赖于颜料 (如 CMYK 模型) 的模型。如图 10.6 所示, 该模型由三个要素组成: 亮度 (L^*) 范围是 0 到 100, a^* 和 b^* 两个颜色通道, 值域由 -128 至 +127。 a^* 包括的颜色是从深绿色 (低亮度值) 到灰色 (中亮度值), 再到亮品红色 (高亮度值); b^* 是从亮蓝色 (低亮度值) 到灰色, (中亮度值) 再到黄色 (高亮度值)。

不难理解这一模型有如下优势:

(1) 不像建模于物理设备输出的色彩空间, $L^*a^*b^*$ 模型致力于人眼颜色感知的均匀性。人们可以通过修改 a^* 和 b^* 分量的输出色阶做精确的颜色平衡, 也可使用修改 L^* 分量来调整亮度对比。在该颜色空间中人们更便于建立评估颜色近似程度的方法, 允许使用数字量方便地表示两种颜色之差。

(2) $L^*a^*b^*$ 空间比显示器、打印机甚至人类视觉的色域都要大, 这样的颜色三要素混合后将产生

具有明亮效果的色彩。 $L^*a^*b^*$ 空间内的很多“颜色”超出了人类视觉的视域，也不能在物理世界中再生。通过色彩管理系统（Color Management System, CMS），这些“颜色”可以通过选择色域内可实现的颜色值来近似。在图像操作的多个步骤之间使用假想色可能是很有用的。 $L^*a^*b^*$ 颜色是 Photoshop 在不同颜色模式之间转换时使用的中间颜色模式。在 Photoshop 的“图像”菜单的“Lab 模式”选项中，软件中去掉了 $L^*a^*b^*$ 中相关的星号*。

(3) 表示为 $L^*a^*b^*$ 的位图比 RGB 或 CMYK 位图获得同样的精度要求更多的每像素数据。对于 8 位/通道的位图，从 RGB 图像到 $L^*a^*b^*$ 之间的来回转换可能有精度损失，但对于现在常见的 16 位/通道甚至 32 位/通道支持，这种损失完全可以避免。

10.2.8 RGB 与 HSV 空间的相互转换

1. 从 RGB 转换到 HSV

在 RGB 空间中的一点 (R, G, B) 可以用下面的一组公式转换成相应的 HSV 空间的 (H, S, V) 坐标：

$$H = \begin{cases} \arccos \left\{ \frac{[(R-G) + (R-B)]/2}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}, & B \leq G \\ 2\pi - \arccos \left\{ \frac{[(R-G) + (R-B)]/2}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}, & B > G \end{cases} \quad (10.2a)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R + G + B)} \quad (10.2b)$$

$$V = \frac{\max(R, G, B)}{255} \quad (10.2c)$$

在求 H 的式 (10.2a) 中，涉及较费计算时间的 \arccos 和开平方运算，因此常用以下的快速算法进行上述的转换，定义

$$\text{Temp1} = \max(R, G, B) \quad (10.3a)$$

$$\text{Temp2} = \min(R, G, B) \quad (10.3b)$$

$$R' = \frac{\text{Temp1} - R}{\text{Temp1} - \text{Temp2}} \quad (10.4a)$$

$$G' = \frac{\text{Temp1} - G}{\text{Temp1} - \text{Temp2}} \quad (10.4b)$$

$$B' = \frac{\text{Temp1} - B}{\text{Temp1} - \text{Temp2}} \quad (10.4c)$$

根据不同的情况，有

$$H_1 = \begin{cases} 5 + B', & R = \text{Temp1} \text{ 且 } G = \text{Temp2} \\ 1 - G', & R = \text{Temp1} \text{ 且 } G \neq \text{Temp2} \\ 1 + R', & G = \text{Temp1} \text{ 且 } B = \text{Temp2} \\ 3 - B', & G = \text{Temp1} \text{ 且 } B \neq \text{Temp2} \\ 3 + G', & B = \text{Temp1} \text{ 且 } R = \text{Temp2} \\ 5 - R', & \text{其他} \end{cases} \quad (10.5)$$

则

$$H = 60H_1 \quad (10.6a)$$

$$S = (\text{Temp1} - \text{Temp2}) / \text{Temp1} \quad (10.6b)$$

$$V = \text{Temp1} / 255 \quad (10.6c)$$

式中, $R, G, B \in [0, 255]$, $H \in [0, 360]$, $S \in [0, 1]$, $V \in [0, 1]$ 。

MATLAB 提供的转换函数为

$$\text{hsvmap} = \text{rgb2hsv}(\text{rgbmap}) \quad \text{或} \quad \text{HVS} = \text{rgb2hsv}(\text{RGB})$$

图 10.7 所示为将一幅 RGB 图像分解成 H 、 S 、 V 分量图的示例。

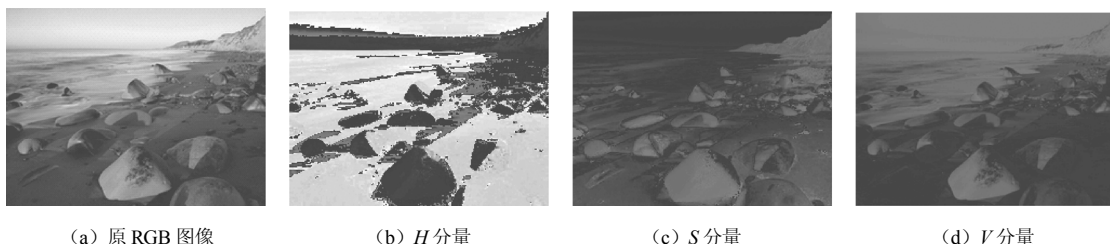


图 10.7 图像的 HSV 分解

2. 从 HSV 转换到 RGB

在给定的 HSV 颜色空间中, (H, V, S) 值定义的一个颜色可以按下面的算法转换成在 RGB 空间中对应的 (R, G, B) 三原色坐标, 其中 R, G, B 为归一化坐标, 变化于 0 到 1 之间:

$$h_i = \lfloor H / 60 \rfloor \bmod 6 \quad (10.7a)$$

$$f = H / 60 - h_i \quad (10.7b)$$

$$p = V \times (1 - S) \quad (10.8a)$$

$$q = V \times (1 - f \cdot S) \quad (10.8b)$$

$$t = V \times [1 - (1 - f) \cdot S] \quad (10.8c)$$

对于每个颜色向量 (r, g, b) ,

$$(R, G, B) = \begin{cases} (v, t, p), & h_i = 0 \\ (q, v, p), & h_i = 1 \\ (p, v, t), & h_i = 2 \\ (p, q, v), & h_i = 3 \\ (t, p, v), & h_i = 4 \\ (v, p, q), & h_i = 5 \end{cases} \quad (10.9)$$

10.2.9 RGB 与 YUV 空间的相互转换

设从一幅图像的某一像素点得到该点在 RGB 坐标空间下的 R 、 G 、 B 值, 则按下式可以得到 Y 、 U 、 V 的值:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.332 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (10.10)$$

图 10.8 所示为是将一幅 RGB 图像的分解成 Y 、 U 、 V 分量图的例子。

从 YUV 转化到 RGB 是上述过程的逆运算, 只要对上述的 3×3 矩阵求逆即可得到。

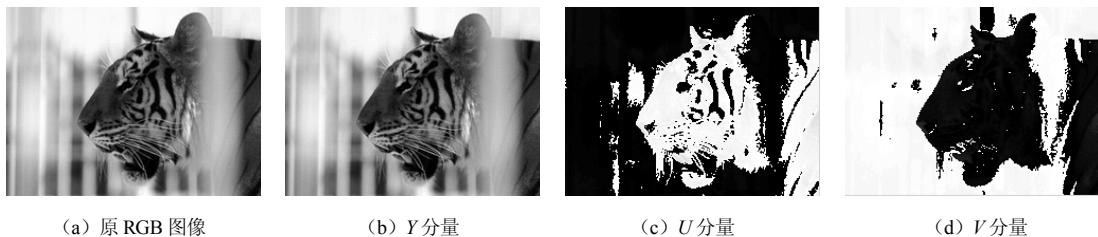


图 10.8 图像的 YUV 分解

10.2.10 RGB 与 HSI 空间的相互转换

1. 从 RGB 空间转换到 HSI 空间

下式定义了以归一化分量形式表示的彩色色调

$$H = \begin{cases} \arccos \left\{ \frac{[(R-G) + (R-B)]/2}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}, & B \leq G \\ 2\pi - \arccos \left\{ \frac{[(R-G) + (R-B)]/2}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}, & B > G \end{cases} \quad (10.11a)$$

颜色饱和度 S 定义为颜色距离 HSI 三角形中心的长度, 其定义如下

$$S = 1 - \frac{3}{R+G+B} \min(R, G, B) \quad (10.11b)$$

三角形仅定义了颜色的色调和饱和度, 并没有给出颜色的强度分量 (I)。对在 $[0, 1]$ 范围内的 R 、 G 、 B 值, 颜色的强度 I 可由以下公式计算

$$I = \frac{1}{3}(R+G+B) \quad (10.11c)$$

注意, 当 $S = 0$ 时对应的是无色的中心点, 这时 H 没有意义, 此时可定义 $H = 0$ 。另外当 $I = 0$ 时对应黑色, H 和 S 也没有意义。

图 10.9 所示为将一幅 RGB 图像分解成 H 、 S 、 I 分量图的例子。

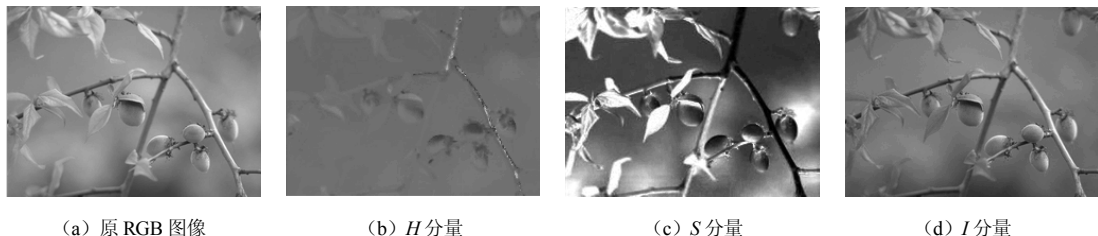


图 10.10 图像的 HSI 分解

2. 从 HIS 空间转换到 RGB 空间

若设 S 、 I 的值在 $[0, 1]$ 之间, R 、 G 、 B 的值也在 $[0, 1]$ 之间, 则从 HSI 到 RGB 的转换公式如下 (分成三段以便利用对称性)。

(1) 当 H 在 $[0, 120^\circ]$ 之间, 即 RG 区域, 先计算 R 、 B , 最后计算 G :

$$R = I \left[1 + \frac{S \cos H}{\cos(60^\circ - H)} \right] \quad (10.12a)$$

$$G = 3I - (R + B) \quad (10.12b)$$

$$B = I(1 - S) \quad (10.12c)$$

(2) 当 H 在 $[120^\circ, 240^\circ]$ 之间, 即 GB 区域:

$$R = I(1 - S) \quad (10.13a)$$

$$G = I \left[1 + \frac{S \cos(H - 120^\circ)}{\cos(180^\circ - H)} \right] \quad (10.13b)$$

$$B = 3I - (R + G) \quad (10.13c)$$

(3) 当 H 在 $[240^\circ, 360^\circ]$ 之间, 即 BR 区域, 先计算 G 、 B , 最后计算 R :

$$G = I(1 - S) \quad (10.14a)$$

$$B = I \left[1 + \frac{S \cos(H - 240^\circ)}{\cos(300^\circ - H)} \right] \quad (10.14b)$$

$$R = 3I - (G + B) \quad (10.14c)$$

10.2.11 真彩色、索引色和灰度图像转换成二值图像

MATLAB 提供的 `im2bw` 函数通过设置亮度阈值 `level` (为相对灰度值, $\text{level} \in [0, 1]$) 可将真彩色图像 RGB、颜色映射表为 `map` 的索引图像 `X`、灰度图像 `I` 转换成二值图像 `BW`。如果是彩色图像, 首先被转换成灰度图像格式, 然后再根据阈值 `level` 转化为二值图像 (相对灰度值高于 `level`, 为 1 (白); 低于 `level`, 为 0 (黑))。阈值 `level` 可以通过调用 `graythresh` 函数自动求得。语法为

`BW = im2bw(I, level)` 或 `BW = im2bw(X, map, level)` 或 `BW = im2bw(RGB, level)`

【例 10.1】 将彩色索引图像 `trees` 按不同灰度阈值转换成二值图像。

【解】 程序如下:

```
load trees
BW1 = im2bw(X, map, 0.2);
BW2 = im2bw(X, map, 0.4);
BW3 = im2bw(X, map, 0.6);
subplot(1, 4, 1); imshow(X, map);
subplot(1, 4, 2); imshow(BW1);
subplot(1, 4, 3); imshow(BW2);
subplot(1, 4, 4); imshow(BW3);
```

运行结果如图 10.10 所示。从转换的效果看, 亮度阈值的选取是十分重要的。阈值过低 (如图 10.10 (b) 所示), 画面出现过白现象, 目标 (图中的树木) 不完整; 阈值过高 (如图 10.10 (d) 所示), 画面出现过黑现象, 目标和背景部分混淆; 取合适的阈值 (如图 10.10 (c) 所示), 目标和背景得到有效的分离, 视觉效果较好。

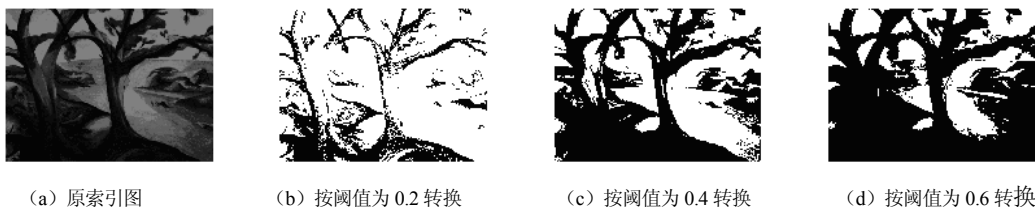


图 10.10 彩色图像转换为二值图像

10.3 颜色空间的量化

10.2.3 节中提到,在许多实用系统中,大量应用与人的视觉特性比较接近 HSV 模型,本节将以 HSV 模型为例,详细描述颜色空间的量化过程。

对于 HSV 色彩空间表示的图像, H 、 S 、 V 任何一个分量都可构成自己的直方图。与 RGB 色彩空间中的色彩直方图相比, HSV 色彩空间各分量的独立性较强,三个直方图相互影响不大,并且主要由 H 色调直方图决定图像的颜色分布。

但是一幅图像色调 H 的范围为 $[0, 360]$, 饱和度 S 的范围为 $[0, 1]$, 亮度 V 的范围为 $[0, 1]$ 。因为人眼对视觉的分辨能力有一定的局限性,因此对整个颜色空间进行适当的量化是必要的,即选择有一定代表性的颜色,丢弃那些无显著视觉意义的信息,从而在解决某些问题时能节约大量的计算空间。如在基于颜色的图像检索系统中采用量化后的颜色空间作为图像的特征空间,可以大大简化后续检索算法的运算量。

如果对 HSV 空间进行适当的量化后再计算直方图,则计算量要小得多。在实际处理中,需要将 HSV 三个分量进行量化以减少特征值的数量。

HSV 色彩空间中颜色特征的量化有两种方式:一种为等间隔量化,另一种为非等间隔量化。在等间隔量化中,依据各自的范围分别按相应的要求均匀等分,然后将之归类,最后以类别作为特征量,比如对色调 H ($0 \sim 360$),可以均分为 6 等份,然后对之归类,即每 60 为一类,这样总共可分为 6 类,即为六维特征量。由于人的视觉对于色调的感知是不均匀的,下面着重介绍根据人的颜色感知特性得到的非等间隔量化提取特征值的方法。

非等间隔量化可依据如下几个方面进行。

(1) 人的视觉分辨能力,将色调 H 空间划分为 8 份,饱和度 S 和亮度 V 空间划分为 3 份。

(2) 根据 H 、 S 、 V 的不同范围和人的主观感知进行色彩量化。

(3) 根据光学理论,颜色与光的波长和频率有关,而不同色光的波长和频率的范围是不同的,因而决定了将色调进行不等间隔量化的准确性。

这样,可以得到以下量化的 H 、 S 、 V 算式

$$H = \begin{cases} 0, & 0 \leq H \leq 20 \text{ 或 } 316 \leq H \leq 360 \\ 1, & 21 \leq H \leq 40 \\ 2, & 41 \leq H \leq 75 \\ 3, & 76 \leq H \leq 155 \\ 4, & 156 \leq H \leq 190 \\ 5, & 191 \leq H \leq 270 \\ 6, & 271 \leq H \leq 295 \\ 7, & 296 \leq H \leq 315 \end{cases} \quad (10.15a)$$

$$S = \begin{cases} 0, & 0 \leq S \leq 0.2 \\ 1, & 0.2 \leq S \leq 0.7 \\ 2, & 0.7 \leq S \leq 1.0 \end{cases} \quad (10.15b)$$

$$V = \begin{cases} 0, & 0 \leq V \leq 0.2 \\ 1, & 0.2 \leq V \leq 0.7 \\ 2, & 0.7 \leq V \leq 1.0 \end{cases} \quad (10.15c)$$

把量化后的 3 个颜色分量合成为一维特征向量

$$I = HQ_SQ_V + SQ_V + V \quad (10.16a)$$

式中, Q_S (分量 S 的量化级数) = 3, Q_V (分量 V 的量化级数) = 3。

$$\text{因此, 式 (10.13) 可表示为} \quad I = 9H + 3S + V \quad (10.16b)$$

这样, 根据式 (10.14), H 、 S 、 V 三个分量可获得 72 柄 (bin) 的一维直方图。

10.4 抖动技术

抖动法 (dithering) 是一种利用仅能重现较少颜色种类的设备来显示含有较多色彩图像的有效方法, 采用能直接显示其色彩的像素模式来替换那些其色彩不能直接显示的像素。抖动技术利用了空间混色原理: 在视觉分辨率内, 人眼能将两种不同颜色的相邻像素融合成第三种颜色。

图 10.11 所示为利用抖动技术在 16 色显示模式下显示全彩色图像的实例。

MATLAB 提供了抖动技术的函数

$$X = \text{dither}(\text{RGB}, \text{map})$$

上述函数通过颜色抖动实现浓淡处理, 把真彩图像 RGB 按指定的调色板 map 转换成索引图像 X, map 的颜色总数不能超过 65536 种。

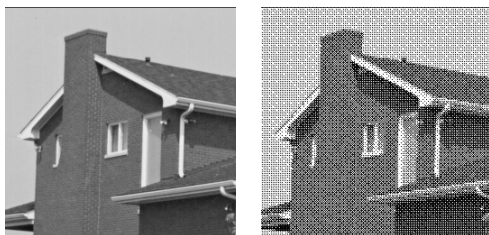
$$\text{BW} = \text{dither}(I)$$

上述函数将灰度图像 I 抖动成二值图像。

$$X = \text{dither}(\text{RGB}, \text{map}, Q_m, Q_e)$$

上述函数利用给定的参数 Q_m 、 Q_e 从真彩色图

像 RGB 中产生索引色图像 X。 Q_m 表示沿各颜色轴反转颜色图的量化位数, Q_e 决定量化误差的位数。如果 $Q_e < Q_m$, 则不进行抖动操作, 返回一个未抖动的索引图像 X。 Q_m 的默认值是 5, Q_e 的默认值是 8。



(a) 原图像 (24 位显示) (b) 抖动图像 (16 色显示)

图 10.11 抖动处理

10.5 假彩色处理

假彩色 (false color) 处理的对象是三基色描绘的自然图像或同一景物的多光谱图像。对自然图像, 假彩色的处理方法之一是将人们所关注的目标物进行线性或非线性映射为与原色不同的假彩色, 即原有的彩色图像变换成给定颜色分布的图像。例如, 绿色草原置成红色、蓝色海洋换成绿色等。这样做的目的是使目标物置于奇特的环境中以引起观察者的注意。方法之二是根据人的眼睛的色觉灵敏度, 重新分配图像成分的颜色。例如, 视网膜中视锥体和视杆体对可见光区的绿色波长比较敏感, 于是可将原来非绿色描述的图像细节或目标物经假彩色处理变成绿色, 以达到提高目标的视觉分辨率的目的。自然图像的假彩色映射可定义为

$$\begin{bmatrix} R_f \\ G_f \\ B_f \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} R_r \\ G_r \\ B_r \end{bmatrix} \quad (10.17)$$

式中, R_r 、 G_r 、 B_r 为原基色分量; R_f 、 G_f 、 B_f 为假彩色三基色分量; T_{ij} 为 ($i, j = 1, 2, 3$) 转移函数。

【例 10.2】 定义一个假彩色映射矩阵, 原图像中的红色变成了蓝色, 绿色变成了红色, 蓝色变成了绿色。

【解】定义以下的基色转换矩阵

$$\begin{bmatrix} R_F \\ G_F \\ B_F \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_f \\ G_f \\ B_f \end{bmatrix}$$

则 $R_F = B_f$, $G_F = R_f$, $B_F = G_f$ 。这样, 原图像中的红色变成了蓝色, 绿色变成了红色, 蓝色变成了绿色。

10.6 彩色图像增强

10.6.1 真彩色增强

“真彩色”(true color)图像是指接近人眼能够分辨的最大颜色数目的彩色图像, 一般指能达到照片质量的 24 位彩色图像。在真彩色增强中, 尽管对 R 、 G 、 B 各分量直接使用对灰度图的增强方法可以增加图像中可视细节的亮度, 但得到的增强图像中的色调有可能没有意义。这是因为在增强图中对应同一个像素的 R 、 G 、 B 这三个分量都发生了变化, 它们的相对数值与原来不同了, 从而导致原图像颜色的较大变化, 且这种变化很难控制。

若将 RGB 图转化为 HSI 图, 亮度分量和色度分量就分开了, 避免了相对数值发生变化。一种采用真彩色增强方法的基本步骤如下:

- (1) 将原始彩色图的 R 、 G 、 B 分量图转化为 H 、 S 、 I 分量图。
- (2) 利用对灰度图增强的方法增强其中的某个分量图。
- (3) 再将结果转换为 R 、 G 、 B 分量图, 以使用彩色显示器显示。

假设增强的是 I 分量, 则不会改变原图的彩色内容, 但增强后的图像看起来可能仍会有些色感不同。这是因为尽管色调和饱和度没有变化, 但亮度分量得到了增强, 整个图像会比原来更亮一些。

对饱和度的增强与对图像亮度的增强有相似的地方。通过对图像中每个像素的饱和度分量乘以一个大于 1 的常数, 可使图像中的彩色更鲜明; 而如果乘以一个小于 1 的常数, 则会使图像的彩色感减少。



图 10.12 图像的饱和度变化

图 10.12 (a) 是一幅原始的彩色图像; 图 10.12 (b) 是增加饱和度分量后得到的结果, 图像的色彩更为饱和, 且有反差增强、边缘清晰的感觉; 图 10.12 (c) 是减小饱和度和度得到的结果, 原来饱和度较低的部分已成为灰色, 整个图像比较平淡。

图 10.13 (a) 是另一幅原始的彩色图像; 图 10.13 (b) 是增加亮度分量后得到的结果, 图像的亮度明显地增强了, 如果亮度分量过大时则图像有点过白; 图 10.13 (c) 是减小亮度得到的结果, 图像的亮度明显降低, 如果亮度分量过小时则会使整个图像偏暗。

下面介绍一种基于小波变换的彩色图像增强算法(见参考文献[20])。为了获得色彩更加鲜艳、细节更加清晰的彩色图像, 根据人眼视觉对亮度和颜色不同的敏感特性, 按照视觉心理选择 HSV 彩色空间, 并在此空间中采用基于小波变换的图像增强算法。



图 10.13 图像的亮度变化

该算法首先进行了自适应亮度动态调节,同时对图像的饱和度分量也进行了处理。和一般的灰度图像增强算法相比,用该算法增强后的彩色图像更加清晰,也更加生动、鲜艳。具体步骤如下:

(1) 将彩色图像变换到 HSV 彩色空间。

(2) 对亮度分量 V 做双正交三次 B 样条函数构造小波变换,在小波域对细节和平滑图像分别采用增强图像的高频信息、调节亮度动态范围算法,然后做小波逆变换。

(3) 对饱和度分量 S 做指数拉伸运算。

(4) 对饱和度、亮度分量 S 、 V 做归一化处理,然后将增强的图像从 HSV 空间转换回 RGB 空间。

10.6.2 伪彩色增强

人眼分辨灰度级的能力较差,一般只有几十级,有时无法从图像的灰度中提取有用信息。而人眼对彩色的分辨能力较强,达几百种甚至上千种。伪彩色(pseudocolor)增强是将灰度或单一波段的图像变换为彩色图像,从而把人眼不能区分的微小的灰度差别显示为明显的色彩差异,更便于解译和提取有用信息。彩色图像中的颜色根据黑白图像的灰度级或其他图像特征(如空间频率成分)人为给定。

伪彩色处理可以用通用计算机完成,也可以用便于实时观察的专用硬设备(伪彩色仪等)来实现。伪彩色处理技术可以在空间域或频率域中实现。常用的伪彩色增强方法有如下 3 种。

1. 利用色彩阶梯变化曲线

这是一种空间域伪彩色处理技术。方法是:把一幅图像按其灰度(L ,共 n 级)不同加上色彩(C)以示区别。如图 10.14 所示,每个灰度区间指定一种颜色 C_i ($i = 1, 2, \dots, n$)。分段的数目及采用的色彩通常由使用者的习惯来决定。

2. 三基色变化法

这也是一种空间域伪彩色处理技术。把图像 $f(x, y)$ 的灰度进行红、绿、蓝三色独立变换,然后把三个变换结果分别单独加到红、绿、蓝三个电子枪上,这样就产生了一幅受变换函数性质所调制的色彩合成图像,如图 10.15 所示。

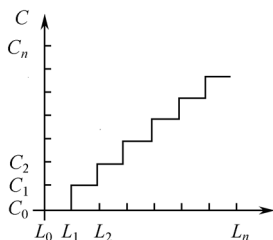


图 10.14 模拟彩色显示的彩色对应变换曲线

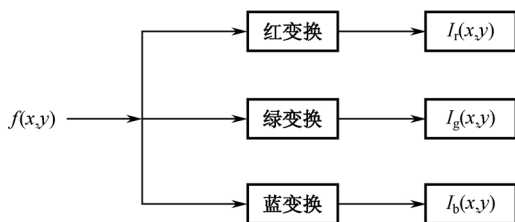


图 10.15 三基色变化法处理框图

3. 频率域伪彩色增强

在频率域中进行伪彩色增强时,先把黑白图像经傅里叶变换到频率域,频率域内三个不同传递特性的滤波器将其分离成三个独立分量,然后对它们进行逆傅里叶变换,便得到三幅代表不同频率分量的单色图像,接着对这三幅图像做进一步的处理(如直方图均衡化),最后将它们作为三基色分量分别加到彩色显示器的红、绿、蓝显示通道,从而实现频率域分段的伪彩色增强。其框图如图 10.16 所示,实例如图 10.17 所示。

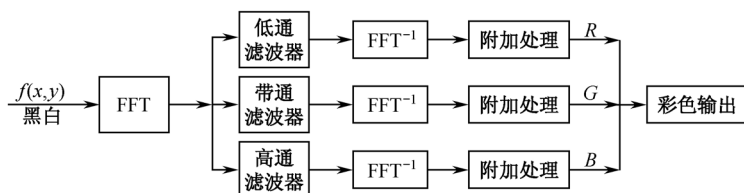
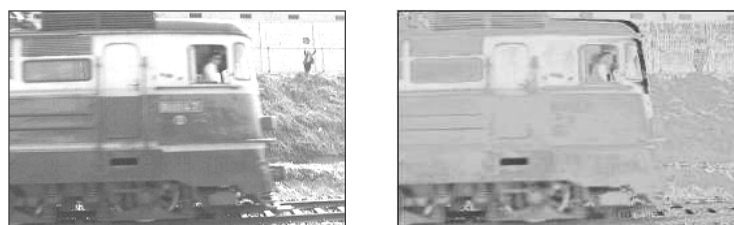


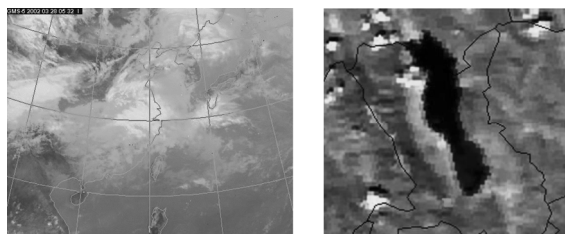
图 10.16 用于伪彩色增强的频域滤波框图



(a) 原图像

(b) 伪彩色增强后的图像

图 10.17 伪彩色图像增强实例



(a) 气象卫星云图

(b) 某地森林火灾的遥感图

图 10.18 伪彩色处理实例

伪彩色处理技术适用于航拍和遥感图片、云图、灾情监测、医学图像的判读等方面, 见图 10.18。例如, 中分辨率成像光谱仪 (Moderate Resolution Imaging Spectroradiometer, MODIS) 是加载于 EOS (Earth Observation System) 卫星上的遥感传感器。MODIS 具有 36 个电磁光谱通道, 是当前世界上新一代“图谱合一”的光学遥感仪器, 有 36 个离散光谱波段, 光谱范围宽, 从 0.4 微米的可见光

到 14.4 微米的热红外全光谱覆盖。为了有效识别、监测和控制污染空气环境的秸秆焚烧现象, 近年来我国某些地方通过 EOS/MODIS 卫星遥感影像解译显示, 以确定秸秆焚烧火点提取准确精度、火点类型、火点影响面积等, 可为环保部门提供环境监测数据, 同时能提高秸秆焚烧预警和日常监控能力。典型的案例利用 MODIS 1B 级数据 (经过定标定位后数据, 该产品是国际标准的 EOS-HDF 格式, 可用商用软件包直接读取) 500 m 分辨率下的第 1、4、3 波段反射率产品, 三个波段分别被赋红、绿和蓝三色叠加进行 RGB 合成伪彩色。基于火点像素亮温特征及其与背景亮温的偏差来提取原秸秆焚烧火点信息。基于 GIS 平台叠加火点信息和地理基础数据生成秸秆焚烧火点分布图。

在质量较高的黑白底片和 X 光片中, 往往有些灰度级相差不大, 却包含着丰富的信息。通常, 将图像中的灰度级转换成不同的彩色, 且分割越细, 彩色越多, 人眼所能提取的信息也越多。现代医用彩色 B 超, 即“彩超 (color Doppler ultrasound)”, 即是伪彩色的典型运用范例。它是在黑白 B 超图像基础上加上多普勒效应原理为基础的伪彩色而形成的。当超声波碰到流向远离探头液体时回声频率会降低, 流向探头的液体会使探头接收的回声信号频率升高。利用伪彩技术加以描述, 使医生能判定超声图像中流动液体的方向及流速的大小和性质, 并将此叠加在二维黑白超声图像上, 形成了彩超图像。图 10.19 所示为颈椎横突附近血管结构彩超图像实例。

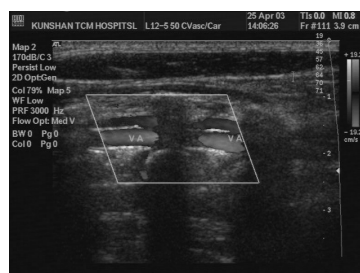


图 10.19 彩超图像实例

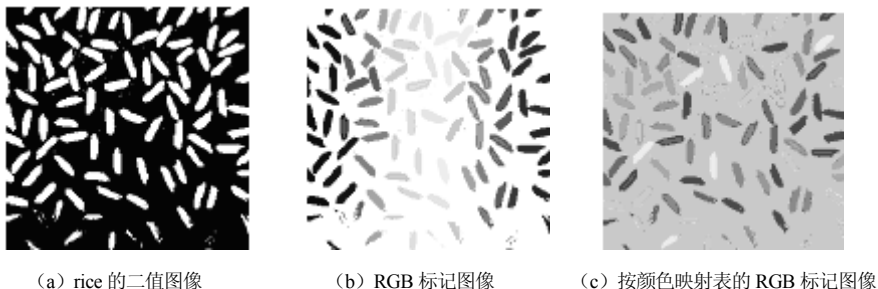
伪彩色还可以用来将标记图像转换为彩色图像，这样我们对目标的观察更加生动。

【例 10.3】 将图像 rice.png 转换标记图像，进一步用转换为 RGB 图像。

【解】 程序如下：

```
I = imread('rice.png');
BW = im2bw(I, graythresh(I));
subplot(131); imshow(BW);
CC = bwconncomp(BW) %返回二值图像的连通分量结构
L = labelmatrix(CC); %由连通分量结构产生标志矩阵
RGB1 = label2rgb(L); %转换标记图像为 RGB 图像
subplot(132); imshow(RGB1);
RGB2 = label2rgb(L, 'spring', 'c', 'shuffle');
%按颜色映射表转换标记图像为 RGB 图像
subplot(133); imshow(RGB2);
```

运行结果如图 10.20 所示。连通分量结构 CC 的信息为：Connectivity: 8; ImageSize: [256 256]; NumObjects: 151; PixelIdxList: {1x151 cell}。由于 151 个米粒用不同的颜色标记，使米粒之间具有更好的区分度，图像细节的观测也更加清晰。该方法同样适用于 SBS 改性沥青材料的显微图像的标记。



(a) rice 的二值图像

(b) RGB 标记图像

(c) 按颜色映射表的 RGB 标记图像

图 10.20 RGB 标记图像

10.7 彩色图像形态学

对于二值图像和灰度图像，数学形态学滤波是一种具有严密理论的非线性几何滤波方法，但从灰度图像形态学向彩色图像形态学的推广仍具有很高的理论和实用价值。

10.7.1 彩色图像形态学的基本方法

彩色图像形态学要解决的主要问题在于彩色图像序结构的建立。二值图像中的“包含”关系和灰度图像的“强度”关系，确立了像素值之间的序结构。但彩色图像的像素是一个多维向量（如 RGB 空间的三维向量），不存在明显的序结构。不同的序结构的建立，便产生了不同的彩色图像形态学方法。

现有的方法可以归纳为两类：分量法和向量法。分量法是利用色彩的空间分解，所形成的每一个子空间（色彩分量）都是灰度图像，从而用灰度形态学分别独立处理各分量；向量法是建立以像素向量为变量的标量函数，以该标量函数值的大小来对向量排序，形成标量函数的灰度图像，再据此应用灰度形态学。为方便起见，本节采用分量法分析彩色图像形态学问题。

在分量法中，灰度形态运算用于彩色图像的每个子空间，在 RGB 彩色空间中，彩色图像 $f(x, y) = [f_R(x, y), f_G(x, y), f_B(x, y)]^T$ 被结构元素 $h(x, y) = [h_R(x, y), h_G(x, y), h_B(x, y)]^T$ 膨胀可以表示如下

$$(f \oplus_C h)(x, y) = [(f_R \oplus h_R)(x, y), (f_G \oplus h_G)(x, y), (f_B \oplus h_B)(x, y)]^T \quad (10.18)$$

式中, \oplus_C 表示彩色形态膨胀, \oplus 表示灰度膨胀。彩色腐蚀、开、闭运算都可以用同样的方法来定义。

以下着重讨论一种基于数学形态学的多尺度彩色图像滤波方法。

10.7.2 基于数学形态学的彩色图像滤波

与灰度图像滤波类似, 彩色图像平滑采用低通滤波, 主要目的是去除图像上的噪声。但在消除或减弱噪声的同时, 对图像细节也有一定的衰减。因此, 图像平滑的直观效果是图像噪声得以衰减, 但同时图像将变得比处理前模糊了, 模糊程度要看对高频成分的衰减程度而定。就同一种平滑方法而言, 去除噪声的效果越好, 图像就越模糊, 因而图像细节损失越多。因此, 在对图像做低通滤波的过程中, 要两者兼顾。

1. 根据相对光视效率曲线确定结构元素

根据红、绿、蓝三种基色的相对光视效率曲线, 人眼对绿色的敏感度最高, 对淡蓝色的敏感程度最低。因此, 我们可以采用多尺度形态学对彩色图像进行滤波。对绿色通道使用小一点的结构元素, 这样可以保留更多的细节信息; 而蓝色通道则可以采用大一点的结构元素, 这样可以滤除更多的噪声; 从而进行加权合并各个通道的滤波结果, 能够比较符合人眼的视觉效果。

2. 多尺度形态学的彩色图像滤波方法

数学形态学作为一种非线性滤波方法, 可克服线性滤波的缺陷, 多尺度形态学滤波就是利用大小不同的结构元素进行图像滤波, 小尺寸的结构元素去噪声能力弱, 但能保存好的边缘细节, 大尺寸的结构元素去除噪声能力强, 但所检测的边缘较粗。因此, 多尺度形态学的彩色图像滤波方法利用大尺度下的抗噪特性抑制噪声, 利用小尺度下的定位特性得到丰富的边缘。可以采用彩色形态开-闭和彩色形态闭-开组合滤波器对彩色图像进行滤波。形态滤波运算定义好了, 另一项重要的工作就是选取多尺度结构元素。

3. 多尺度结构元素的选取

假定一个结构元素序列 $\{g_i \mid i = 1, 2, \dots, n\}$ 具有相同形状且尺寸随 i 增加而单调增大, 则定义序列 $\{g_i\}$ 为一个多尺度序列。如果合理地调节结构元素尺度的大小, 就可以抽取图像不同尺度上的特征, 能够有效地克服噪声并且得到满意的边缘信息。一般取基本有限结构元 g_1 (取为 3×3 十字形), 且 g_i 的大小随 i 增加而单调增大, 即为 $(2i+1) \times (2i+1)$ 个像素 (一般 i 取 $1 \sim 3$ 即可), 如图 10.21 所示, 其中 “★” 构成结构元素。

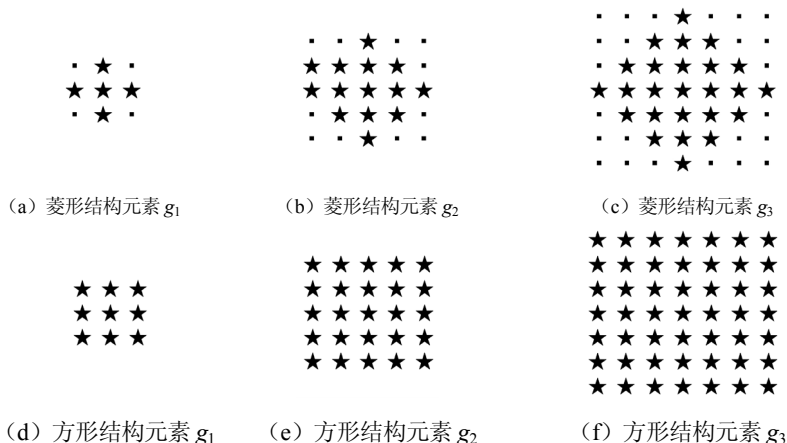


图 10.21 多尺度形态结构元素

4. 彩色图像形态学滤波器的滤波效果分析

为了分析和评价上面提出的形态学滤波器的滤波效果,采用 24 位真彩色 Lena 图像作为测试图像,尺寸为 512×512 ,并加入了 5% 的椒盐噪声,红、绿、蓝三通道分别选用图 10.21 中的方形结构元素 g_1 、 g_2 、 g_3 ,单一结构元素选取菱形 g_1 ,将计算结果与其他滤波器相比较,比较结果如图 10.22 所示。中值滤波器对椒盐噪声的滤波效果最好,多结构元素滤波效果和中值滤波器性能差不多,单一结构的滤波性能差一些,维纳滤波器和均值滤波器滤除椒盐噪声的效果更差。原始噪声图像经过均值滤波后,噪声得到了抑制,图像也得到了平滑,但同时也使图像边缘变得模糊了。同样,由于使用的结构元素的不同,滤波效果也有差别,测试结果如图 10.23 所示。可见,菱形 g_1 结构元素滤波效果最佳,方形 g_1 结构元素滤波效果次之,而方形 g_2 结构元素滤波效果最差。



图 10.22 不同滤波器的滤波结果比较



图 10.23 不同结构元素滤波结果比较

10.8 实验：彩色图像处理

1. 实验目的

- (1) 了解颜色空间的基本概念及转换原理。

(2) 能利用 MATLAB 工具箱和本章提供的有关公式, 在 RGB、YIQ、HSV 等颜色空间中读取、显示彩色图像。

(3) 学习 Photoshop 颜色的调整。

(4) 学习 Photoshop 调色板的使用。

2. 实验主要仪器设备

(1) 台式机或笔记本电脑。

(2) MATLAB (安装了图像处理工具箱)、Photoshop。

(3) 彩色图像文件。

3. 实验原理

(1) 利用 MATLAB 图像处理工具箱读、写和显示图像文件。

(2) 颜色空间的基本概念。

① 颜色空间包括某个颜色域的所有颜色。实验中采用的颜色模型有 NTSC、HSV 和 YCbCr 模型等。各种颜色空间在图像处理中各有优势, 且可以相互转换。

② 在 NTSC 格式中, 图像由三个分量表示: 亮度用 Y 表示; 色度用 I 表示; 饱和度用 Q 表示。第一个分量亮度 Y 表示灰度信息, 后两个分量分别表示色度信息和饱和度信息。RGB 空间向 YIQ 空间转换方法为

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (10.19)$$

③ HSV 模型常用于选择颜色, 它是面向用户的一种复合主观感觉的色彩模型, 比 RGB 模型更接近人们对颜色的感知。具体转换参考式 (10.2) ~ 式 (10.9)。

④ 在 YCbCr 模型中, Y 为亮度, C_b 和 C_r 共同描述图像的色调, 其中 C_b 和 C_r 分别为蓝色分量和红色分量相对于参考值的坐标。RGB 空间向 YCbCr 空间转换的方法为

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (10.20)$$

(3) Photoshop 颜色的调整, 包括分别调整彩色图像的有关参数 (亮度、对比度、灰度系数等, 图像增强实验中已完成) 和变色 (Variations)。如图 10.24 所示, 图框上的 6 个图框表示图像可以变得更 Green (绿)、Yellow (黄)、Red (红)、Magenta (品红)、Blue (蓝)、Cyan (青), 需要哪一种, 只需单击相应的框。右边的 3 个图框分别表示较亮 (Lighter)、当前挑选和较暗 (Darker)。



图 10.24 颜色的变色调整

(4) Photoshop 调色板的使用。有时需要把一幅真彩色的图像文件转换为 256 色。Photoshop 可以计算出图像所需的最佳调色板, 也允许我们为图像指定一个调色板。用指定的调色板来转换 256 色图像这一功能非常有用, 在制作多媒体软件时, 我们经常需要使很多张图使用同一个调色板, 这样才能保证在 256 色显示模式下正确显示。

4. 实验内容

(1) 利用 MATLAB 图像处理工具箱中现有的函数在不同颜色空间中读取、显示彩色图像。

- (2) 根据转换公式, 自己编制其他颜色空间转换函数, 在该空间中读取、显示彩色图像。
- (3) Photoshop 颜色的调整。
- (4) Photoshop 调色板的使用。

5. 实验步骤

- (1) 利用 MATLAB “读图像文件 I/O” 函数读入图像。
- (2) 参考附录 B “常用 MATLAB 图像处理工具箱函数” 完成以下操作: 利用 `rgb2ntsc`、`rgb2hsv`、`rgb2ycbcr` 等函数转换颜色空间, 显示不同颜色空间下的彩色图像, 并显示出各分量图像。例如, 在 YCbCr 空间中, 通过编程分别显示并观察 Y、Cb、Cr 三分量的图像。
- (3) 根据其他颜色空间转换公式, 编制颜色空间转换函数, 在该空间中读取、显示彩色图像。例如, 在 HSI 空间, 参考式 (10.11) ~ 式 (10.14)。
- (4) 使用 Photoshop 调色板。

6. 实验报告要求

- (1) 利用 MATLAB 图像处理工具箱、Photoshop 读、写和显示彩色图像文件。
- (2) 了解几种常用颜色空间的概念。
- (3) 记录读取变换处理后的图像。
- (4) 列出上述 MATLAB 图像处理的程序或 Photoshop 界面。
- (5) 回答思考题。
- (6) 心得和体会。

7. 预习要求

- (1) 了解 MATLAB 图像处理工具箱中所提供的颜色空间转换函数的功能。
- (2) 了解图像颜色空间转换的基本概念和公式。
- (3) Photoshop 颜色的调整、调色板的使用。

8. 思考题

- (1) 比较不同颜色空间的差别。
- (2) Photoshop 颜色调整的其他功能的实现原理、操作步骤及效果。如色彩平衡、反相、色调匀化等。

本章小结

人的颜色视觉的产生是个复杂的过程, 它包含色度学和光度学知识。不同的彩色图像处理场合需要选择合适的颜色空间, 常用的颜色空间有 RGB、HSV、HSI、YUV 等, 这些颜色空间各有特点。人眼对视觉的分辨能力有一定的局限性, 因此对整个颜色空间进行适当的量化是必要的。HSV 色彩空间中颜色特征的量化有两种方式: 一种为等间隔量化, 另一种为非等间隔量化。

抖动算法是利用仅能显示少数彩色设备来显示含有丰富色彩信息图像的一类有效方法。

假彩色处理的对象是三基色描绘的自然图像或同一景物的多光谱图像。

伪彩色增强是将一个波段或单一的黑白图像变换为彩色图像, 从而把人眼不能区分的微小的灰度差别显示为明显的色彩差异, 更便于解译和提取有用信息。常用的伪彩色图像处理方法有三种: 利用色彩阶梯变化曲线、三基色变换法和频率域伪彩色增强。

真彩色增强通过对彩色图像的不同分量分别进行映射来改变原来彩色图像的视觉效果。

思考题与习题

- 10.1 如何表示图像中一点的彩色值？颜色模型起什么作用？
- 10.2 色调、色饱和度和亮度的定义是什么？在表征图像中一点的颜色时，各起什么作用？
- 10.3 解释在同一个显示器上 Windows 中打开照片和 Photoshop 中打开照片颜色会出现不一致的现象。
- 10.4 为什么有时需要将一种颜色数据表示形式转换为另一种形式？如何由 RGB 数值计算 HSV 数值？
- 10.5 当 $R=0$, $G=0$, $B=1$ 时，在 HSI 空间求 H 和 S 值。
- 10.6 什么是彩色的减性模型和加性模型？哪一种模型更适合用于显示和打印图片场合？
- 10.7 哪个颜色空间最接近人的视觉系统的特点？
- 10.8 完成 RGB 空间与 HSI 空间相互转换的程序。
- 10.9 将彩色索引图像 trees 按不同灰度 graythresh 函数计算的阈值转换成二值图像。
- 10.10 为什么在某些场合下要进行彩色量化？彩色图像的量化依据是什么？
- 10.11 编程实现 HSV 空间彩色图像颜色特征的非等间隔量化算法。
- 10.12 抖动技术是如何利用只能显示较少颜色的设备重现含有丰富色彩图像的？
- 10.13 用 dither 函数将灰度图像 cameraman.tif 抖动成二值图像。
- 10.14 编写一个将原图像中的红色变成了蓝色、绿色变成了红色和蓝色变成了绿色的 MATLAB 程序。
- 10.15 讨论伪彩色在云图绘制和显示等气象领域的应用。
- 10.16 讨论假彩色和伪彩色的差异。
- 10.17 讨论彩色图像增强与灰度图像增强的关系。

第 11 章 数字图像处理的应用

➡ 内容提要

数字图像处理的应用十分广泛，本章介绍现代数字图像处理的部分应用实例，包括图像处理在数字水印上的应用、基于数学形态学的图像颗粒度分析系统、数字化医院中的图像存档与通信系统 (PACS)、基于多分辨率分析的图像融合方法和基于 MATLAB GUI 图像处理软件的开发简介等，本章实验要求读者应用已学的图像处理知识，独立设计具有一定功能的基于 GUIDE 图像处理软件。

➡ 知识要点

- 数字图像水印系统。
- 图像通信系统。
- 基于内容的图像检索系统。
- 图像分析系统。
- 图像融合技术。
- 基于 MATLAB GUI 图像处理软件的开发。

➡ 教学建议

- 本章教学安排 4 学时，作为选讲或自学内容，部分内容可以作为讲座、课程设计或毕业设计的素材。
- 学习本章需要前面各章的基础，同时也需对相关的理论和技术进行扩展，更偏重于应用型、系统性和前瞻性。
- 本章的目的旨在让读者了解图像处理的具体应用，为读者结合实际要求学习、使用和设计图像处理系统打下基础。
- 通过基于 MATLAB GUI 图像处理软件的开发技术，读者可以逐步完成从理论与技术阶段的原理学习到工程实践的系统设计和开发应用的飞跃。这一过程的磨练很重要，不可或缺，读者会从中得到许多实际工程经验积累的新体验。

11.1 图像处理在数字水印上的应用

11.1.1 概述

随着信息时代的到来,特别是互联网的普及,信息的安全保护问题日益突出。当前的信息安全技术基本上以密码学理论为基础,无论是采用传统的密钥系统还是公钥系统,其保护方式都是控制文件的存取,即将文件加密成密文,使非法用户不能解读。但随着计算机处理能力的快速提高,这种通过密钥来提高系统保密性的方法变得越来越不安全。

另一方面,多媒体技术已被广泛应用,需要进行加密、认证和版权保护的声像数据也越来越多。如果对数字化的声像数据也采用密码加密方式,则其本身的数字信号属性就会被忽略。于是,许多研究人员尝试用各种信号处理方法对声像数据进行隐藏加密,并将该技术用于制作多媒体的“数字水印”(digital watermark)。数字水印是信息隐藏(information hiding)技术的一个重要研究方向。

数字水印技术是指用信号处理的方法在数字化的多媒体数据中嵌入隐蔽的标记,这种标记通常是不可见的,只有通过专用的检测器或阅读器才能提取。

11.1.2 数字水印的衡量标准

水印的评价一般是凭视觉主观判断其质量优劣。最常用的衡量水印方案的标准有两个:

- (1) 不可见性(隐蔽性):在数字作品中嵌入数字水印不会引起明显的降质,并且不易被察觉。
- (2) 鲁棒性(robustness):鲁棒性是指在经历多种无意或有意的信号处理过程后,数字水印仍能保持完整性或仍能被准确鉴别。可能遇到的信号处理过程包括施加信道噪声、滤波、数/模与模/数转换、重采样、剪切、位移、尺度变化以及有损压缩编码等。

鲁棒性可以通过计算提取出的水印与原水印的归一化相关系数 NC 和峰值信噪比 PSNR 来进行评价。

11.1.3 数字水印的分类

我们可将数字水印按以下四种方式进行分类:

- (1) 按水印的特性,可以将数字水印分为鲁棒数字水印和脆弱数字水印两类。鲁棒数字水印要求嵌入的水印能够经受各种常用的编辑处理;脆弱数字水印则需要对信号的改动很敏感,使人们能根据脆弱水印的状态判断数据是否被篡改过。

- (2) 按水印的检测过程,可以将数字水印划分为明文水印和盲水印。明文水印在检测过程中需要原始数据,而盲水印的检测只需要密钥,不需要原始数据。一般来说,明文水印的鲁棒性比较强。

- (3) 按数字水印的内容,可以将水印划分为有意义水印和无意义水印。有意义水印是指水印本身也是某个数字图像(如商标图像)或数字音频片段的编码;无意义水印则只对应于一个序列号,通过统计决策来确定信号中是否含有水印。

- (4) 按数字水印的隐藏位置,可以将其划分为空间域数字水印、频率域数字水印、空间-频率域数字水印等。空间域方法的鲁棒性较差,水印信号容易丢失,因此目前的研究方法主要集中在频率域。

11.1.4 实现数字水印的一般步骤

可大致将实现一个数字水印系统的过程分为三个阶段：嵌入过程、传播过程和抽取过程。其中嵌入和抽取是相互对应的，即不同的嵌入方法对应着不同的抽取方法。水印方案的提出要充分考虑到数字产品在传播过程中会受到怎样的干扰，这些干扰可能是天然的，比如信道噪声；也可能是人为的，比如恶意地篡改数字产品。这些都需要被充分估计，并设计出能够抵抗这些干扰的水印方案。

在嵌入之前，要对嵌入载体做一些预处理或变换，同样，抽取水印也需要做相同的工作，所以一般的数字水印的实现流程如图 11.1 所示。

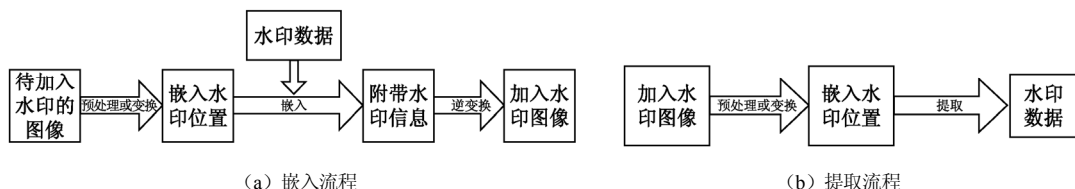


图 11.1 水印的嵌入和提取流程

11.1.5 图像水印举例

1. LSB 方法

这是一种最简单的嵌入水印的办法。事实上，任何一幅图片都具备一定的容噪性，这表现在像素数据的最低有效位（Least Significant Bit, LSB）对人眼的视觉影响很小，秘密信息就隐藏在图像每一个像素的最低位或次低位，实现其不可见性。以 256 色灰度图像为例，每个像素值占 8 bit，其第 8 位就是最低有效位。

具体的 LSB 方法就是调整原始载体信息的最低几位来隐藏信息，最方便的是采用直接改变图像中像素的最后一位使之和秘密信息相同。检测的时候只要提取含水印图像像素的最低位即可。为了说明问题，我们把水印分别嵌入到图像像素的不同位，从图 11.2 可以明显地看出，越低位嵌入，人眼越难识别。



图 11.2 用 LSB 方法嵌入水印

2. 在 DCT 域嵌入水印

先计算图像的离散余弦变换 (DCT), 然后将水印叠加到 DCT 域中幅值最大的前 k 个系数上 (不包括直流分量), 通常为图像的低频分量。若 DCT 系数的前 k 个最大分量表示为 $D = \{d_i\}$ ($i = 1, 2, \dots, k$), 水印服从高斯分布的随机实数序列 $W = \{w_i\}$ ($i = 1, 2, \dots, k$), 那么水印的嵌入算法为

$$D_i = d_i(1 + \alpha w_i) \quad (11.1)$$

其中, 常数 α 为尺度因子, 控制水印嵌入的强度。然后用新的系数做逆变换得到水印图像 I_w 。解码函数则分别计算原始图像 I 和水印图像 I_w 的离散余弦变换, 并相减得到水印估计 W^* , 再和原始水印做相关检验以确定水印的存在与否。一般地, 如果水印存在, 相关系数应该很大, 没有水印时则很小, 接近于 0。由于水印是随机高斯噪声, 嵌入前后的图像人眼难以区分。

根据水印的分类, 第一种方案的检测不需要原始图像, 所以是盲水印, 同时又是有意义的水印和空间域水印。但是, 这种水印是非常脆弱的, 简单的滤波就会使提取的水印信息面目全非, 所以它又是脆弱水印。

同理, 第二种方案嵌入的是非盲、鲁棒和频率域的水印。需要说明的是, 由于其水印信号是随机的高斯序列, 我们只是用来判断图像中是否含有水印, 更多的信息无法得到, 所以是无意义水印。

3. 在 DWT 域嵌入水印

这里介绍一种将彩色数字水印嵌入到原始彩色数字图像中的算法。该算法将水印多次嵌入到离散小波变换 (DWT) 后相应的频段来增强鲁棒性; 利用了 HVS 特性, 并通过实验获得水印不同彩色分量 (R, G, B) 的加权系数, 从而使嵌入水印后的图像无主观视觉上的失真。通过实验验证, 该算法在对嵌入水印后的图像进行 JPEG 有损压缩、剪切、不规则色块污染、添加各种噪声处理后, 提取出的水印表现出良好的鲁棒性, 且主观视觉失真较小。

图 11.3 所示为小波域水印嵌入的一般框图。

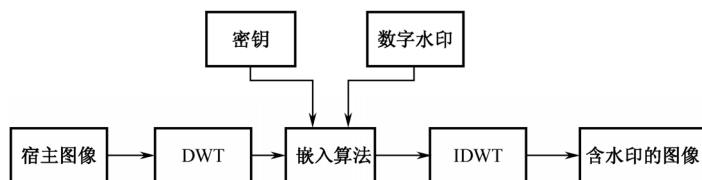


图 11.3 小波域水印嵌入的一般框图

图 11.4 所示为小波域水印检测的一般框图。

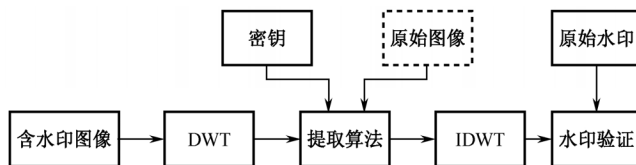
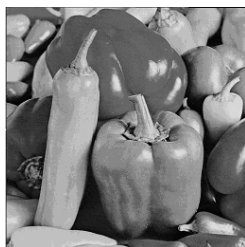
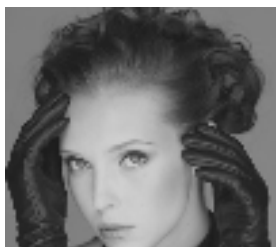


图 11.4 小波域水印检测的一般框图

图 11.5 所示为提取水印的一个示例。



(a) 含水印的 Pepper 图像




(b) 提取出的水印

图 11.5 水印的提取

对于各种攻击情况，对算法性能进行测试。如 JPEG 压缩攻击、高斯和椒盐噪声攻击、剪切攻击等。表 11.1 以受剪切攻击为例，列出了水印的检测结果，可见在受到较严重的剪切攻击时，该算法仍能有效地检测出水印。

表 11.1 剪切攻击结果

剪切后含水印图像				
提取水印				
PSNR	18.65	20.95	23.18	20.70

11.2 基于数学形态学的图像颗粒度分析系统

11.2.1 概述

在实际问题中，一张细胞图片中可能包含多个彼此分离的细胞；金相图片中可能有多个气泡或颗粒；遥感图片中可能有许多不同的目标。在建筑工程中，通过绘制混凝土的颗粒度（granularity）分布函数，可以在一定程度上得出混凝土的质量是否达到预期的标准。医学图像中通常存在着一些由电子显微镜等图像获取设备造成的可见颗粒噪声，它们会降低图像的质量。为了处理显微寄生虫卵图像中的目标颗粒问题，本系统采用了一种基于数学形态学的去噪及颗粒统计算法。首先，选用正方形和圆形两种结构元素进行开运算进行去噪处理得到了不同的效果。其次，对于同一种结构元素，其去噪效果随结构元素半径（radius）的改变而产生明显的差异。然后，再分别利用这些结构元素进行颗粒度的检测与计算，取得了较好的检测效果。

11.2.2 求图像中目标的面积和颗粒度

1. 面积的求解

面积的求解方法大致可以分为如下几种。

(1) 几何方法

假设在将连续图像离散化时所使用的网格为正方形，而且相邻两网格的中心点间距离为 c ，则每

个网格的面积为 c^2 。另一方面,若网格为正六边形且相邻两网格的中心点距离为 c ,则每个网格面积为 $(\sqrt{3}/2)c^2$ 。

现假定在一幅数字化二值图像中对应于目标图像的部分为 Y ,用 $N(1)$ 表示 Y 中像点(灰度为1)的个数。那么,当使用正方形网格时,与 Y 相应的连续目标图像的面积

$$A(Y)=c^2N(1) \quad (11.2)$$

当在使用正六边形网格时,与 Y 相应的连续目标图像的面积

$$A(Y)=\frac{\sqrt{3}}{2}c^2N(1) \quad (11.3)$$

由此可见,在由离散图像估计连续图像面积时应考虑到网格形状。选取网格网状时应注意保持离散化后图像的分辨率。为了做到这一点,两个格之间的距离必须不大于连续目标图像的最小曲率半径。

如果事先已知一连续图像的最小曲率半径为 r_0 ,则在使用正方形网格对它实行离散化时,相邻(包括在对角线方向上有一共同点的情形)两网格点间的距离不能大于 r_0 ,即网格边长不大于 $r_0/\sqrt{2}$,而在使用正六边形网格时,点间距离不得大于 r_0 ,即边长不大于 $r_0/\sqrt{3}$ 。

(2) 运用格林公式

用边界坐标计算面积,根据格林定理,在 x - y 平面中的一个封闭曲线包围的面积由其轮廓积分给定,即

$$A=\frac{1}{2}\oint(xdy-ydx) \quad (11.4)$$

式中,积分沿着该闭合曲线进行。将具有 N_b 个像素的封闭曲线离散化成 (x_i, y_i) ($i=1,2,\dots,N_b$),得

$$A=\frac{1}{2}\sum_{i=1}^{N_b}[x_i(y_{i+1}-y_i)-y_i(x_{i+1}-x_i)]=\frac{1}{2}\sum_{i=1}^{N_b}[x_iy_{i+1}-x_{i+1}y_i] \quad (11.5)$$

(3) 用边界行程码或链码

由各种封闭边界区域的描述来计算面积,可分为如下情况:

① 已知区域的行程编码,只需把值为1的行程长度相加,即为区域面积。

② 若给定封闭边界的某种表示,则相应连通区域的面积应为区域外边界包围的面积与内边界包围的面积(孔的面积)之差。

2. 三种面积计算方法的比较

几何方法从图像本身离散化的网格分布出发,使得求解复杂度随图像离散化的网格数的增加而增大,不利于编程。运用格林公式有利于编程,在求解前还应考虑图像中目标水平和垂直方向的比例因子,计算复杂度不大。用边界行程码或链码虽然比较简单,但边界行程码或链码依情况而定,程序复杂度与运用格林公式相当。

3. 颗粒度的求解

对图像进行预处理后,可提取图像中相关的特征信息。本系统主要对图像中颗粒度进行检测,用“探针”结构元素与图像作形态学运算。求图像中目标的颗粒度主要分以下两步。

(1) 颗粒的检出

在研究图像的颗粒分布特征时,首先需要从图像中检出那些具有实际意义的颗粒或子图像 Y ,然后消除噪声点。对于一幅图像 X ,选取结构元素 λB (B 是半径为1的结构元素, λ 是半径)进行腐蚀

运算, 去掉半径小于 λ 的噪声点, 再进行膨胀运算, 使去噪后的图像恢复原状。综合起来, 对 X 进行了一次开运算, 只要适当地选取结构元素并利用开变换, 便可达到检出颗粒的目的。

(2) 颗粒分布函数

在消除了噪声点后, 就是对 X 中全体颗粒 Y 的分布状况作出一个定量的描述。由此定义以下的“颗粒分布函数”

$$f(\lambda) = \frac{A(X \circ \lambda B) - A(X \circ (\lambda+1)B)}{A(\lambda B)} \quad (11.6)$$

$X \circ \lambda B$ 表示用 λB 对 X 进行开变换后所保留下来的全体颗粒, $A(X \circ \lambda B)$ 是这些颗粒的总面积, $A(X \circ (\lambda+1)B)$ 是用 $(\lambda+1)B$ 进行筛选后所剩下的颗粒的总面积, 而 $A(\lambda B)$ 是 λB 的面积。通过编程绘制“颗粒分布函数”, 显然, 结构元素的半径越大, 则被保留的颗粒越少, 因此恒有 $f(\lambda) \geq 0$ 。若 X 中的颗粒普遍较大, 则筛选后所保留的数量也较多, 这时 $f(\lambda)$ 的值也较大。

通过这个函数表达式, 可知如果图像中的颗粒形状与结构元素具有较高的匹配度, 那么图像颗粒度 $f(\lambda)$ 值将变得较大。故在检测图像颗粒度的时候, 选取了合适的结构元素, 才可能得到较好的实验结果。

11.2.3 实验结果与分析

在实际问题中, 很多图像是互不联结而呈颗粒状分布的。本系统选择合适的结构元素进行滤波并绘制出面积函数和颗粒度分布函数。

1. 结构元素的选取

选取了正方形和圆形两种结构元素, 如图 11.6 所示。圆形结构元素用半径 r 表示, 而正方形结构元素的半径大小指中心像素点到各边缘像素点间的像素的数目, 用 d 来表示。

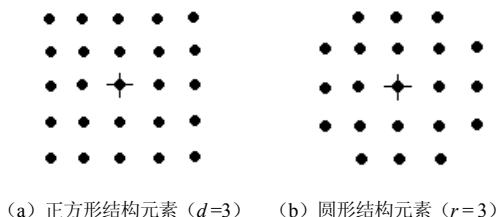


图 11.6 两种结构元素

2. 去噪预处理

传统的滤波方法虽然很好地去除了噪声, 但它同时去除了一部分图像信息, 如图像的边缘信息, 造成图像边缘的模糊。基于形态学的去噪, 它利用了结构元素对图像进行处理, 以提取有用的信息或者除去噪声, 并能够很好地保留图像的边缘信息。灰度形态学在去除图像噪声方面与二值形态学相似, 但灰度形态学处理的对象是灰度图像, 所以它具有独特的特点。去噪的算法主要由以下几个步骤完成:

① 读入原始图像, 将彩色图像转换为灰度图像 (见图 11.7 (a))。

② 加上椒盐噪声 (见图 11.7 (b))。

③ 选取合适的结构元素 (见图 11.6)。

④ 对含有噪声的图像进行开运算, 去除亮斑点; 对得到的图像再进行闭运算, 消除图像中的暗斑点, 即滤除了图像中相应的噪声 (见图 11.7 (c) 和图 11.7 (d))。

仿真中以一幅寄生虫医学图像 (256×256) 为例。分别采用正方形结构元素 ($d=3$) 与圆形结构元素 ($r=3$) 进行去噪。仿真结果表明, 结构元素的选取对去噪结果产生很大的影响, 图 11.7 (c) 很好地消除了噪声, 同时较好地保留了原图像的信息。而图 11.7 (d) 虽然也抑制了噪声的干扰, 但原图像边缘被模糊化了, 改变了原图像特征信息, 影响对有用信息的提取。通过图 11.7 (c) 与图 11.7 (e) (或图 11.7 (d) 与图 11.7 (f)) 相比较, 可知对于同一种结构元素, 选取半径 (d 或 r) 为 4 的结构元素得到的去噪效果更差, 图像变得更模糊, 而且出现了干扰斑点。

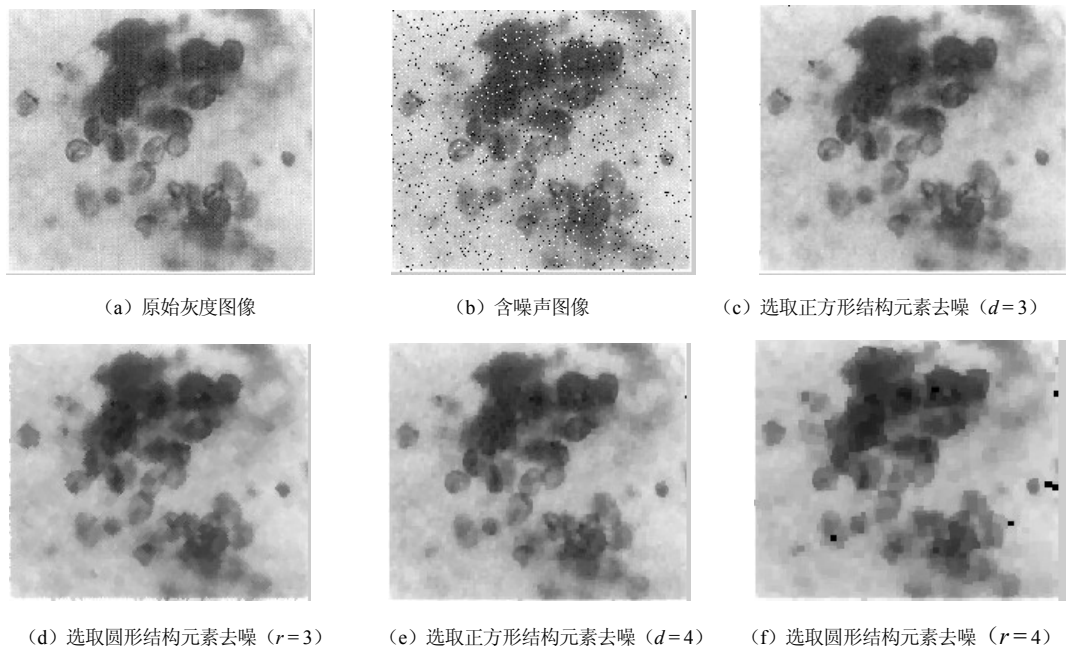


图 11.7 显微图像的形态学去噪处理

3. 计算和绘制颗粒度分布函数

颗粒分布函数值总体上随着结构元素的变大而变小。而且, 选择不同大小的结构元素所能检测的颗粒度大小范围不同, 并且所保留各种大小的颗粒数量也有很大的差别。相比于正方形结构元素, 由于圆形结构元素比较接近于被检测图像中颗粒的形状, 利用圆形结构元素能更好地检测出图像中颗粒各大小颗粒的数量。

11.2.4 小结

结构元素对许多实际问题的形态运算结果有决定性的作用, 因此, 需结合应用背景和期望结果合理选择结构元素的形状和大小。颗粒度分布函数在一定程度上可以反映图像中的颗粒分布状况。当面积函数接近于一定值, 说明图像颗粒半径普遍小于这个值。结构元素的半径越大, 经过开运算后留下的颗粒越少, 图像面积会迅速减少而接近于零。根据图像的面积函数变化情况可以帮助我们选择更合适的结构元素的半径。

11.3 基于内容的图像检索 (CBIR)

11.3.1 概述

随着互联网技术、计算机技术、信号处理技术等不断发展, 每天都有大量的图像数据需要处理, 图像数据的应用领域已涉及科学技术和日常生活的各个方面。如在气象、遥感领域, 人们要处理海量的气象图片和遥感图片。在医疗卫生领域, 由于医学图像的数字化存档和通信系统 (PACS) 已经越来越为各医疗机构所重视, 大中医院普遍采用各种各样的医学影像设备帮助医生进行诊断治疗, 医学影像设备 (如 CT、MRI、X 光、内窥镜、超声波设备、心电图机等) 产生的图像数据很大部分已实现了

数字化管理。在日常生活中,数码相机、扫描仪、数字摄像机等信息电器(IA)已进入寻常百姓家中,使得我们电脑硬盘或光盘里可能存储了成千上万张人物、风景图片。

医学图像的大量出现给图像的查询工作带来了极大的不便,所以基于内容的医学图像检索就显得很重要。举一个典型的例子:某位医生在诊断时遇到疑难病症,他需要找到图像库中相似的图像,以便参考已有的确诊病史来帮助诊断,这时就需要进行基于内容的医学图像检索。由于医学图像的主要使用者是医生,医生可以利用图像生成时间、患者姓名、诊断医生姓名、图像设备等标注进行查询。但医生能够利用颜色、纹理、形状等特征对医学图像进行精确的解读,他们提供的图像内容信息较为专业和客观,这就为基于语义的医学图像查询提供了可能。

由此可见,图像数据与日常生活的关系越来越密切,这就要求我们要能对图像数据进行科学管理,与此同时,人们感到要在海量的图像中寻找到自己所需的图像变得越来越困难,于是人们开始对图像检索领域进行研究。虽然在过去的十多年里,基于内容的图像检索在理论研究和系统开发方面都取得了令人瞩目的成就,但仍然存在许多挑战性的问题值得研究者关注,所以这个领域的研究方兴未艾。

11.3.2 基于内容图像检索的发展

从 20 世纪 70 年代起,人们开始对图像检索进行研究。最初人们希望能像传统数据库那样对图像进行基于文本的查询,即利用传统数据库技术去管理图像。这种方法首先要对图像进行文本标注,其次对标注图像进行基于文本的查询。基于文本检索的图像数据库通过赋予图像一组特征数据来表达图像的内容,从而将对图像本身的检索转化为基于文本的检索,它利用了自然语言的优势,间接地实现了对图像的索引。但这种方法不可避免地存在下列问题:

(1) 人工标注要受到人的主观意识限制,而图像内容又是如此丰富多彩,不同的人对同一幅图像的理解很难相同,这就有可能在检索时出现歧义性。

(2) 大量的图像进行人工标注时要花费大量精力,特别是对大型的、不断更新的图像数据库,这项工作更是令人乏味。

(3) 没有直接基于图像的视觉特征,没有充分利用图像本身所包含的特征如颜色、纹理、形状等,也没有充分利用计算机对数字图像的处理能力。

出现这些问题的根本原因是,目前的应用程序无法从图像中抽取与语义相关的文本,所以人们无法使用现有的高效文本检索系统来查询图像。

进入 20 世纪 90 年代后,由于互联网和新的数字图像传感器技术的迅速发展,使得各种图像大量产生。显然,基于文本的图像检索系统越来越不能满足应用的需要,对日益膨胀的视觉信息进行有效的管理已经成为迫切需要解决的问题,因而基于内容的图像检索(Content-based Image Retrieval, CBIR)技术应运而生。需求是基于内容图像检索技术出现的内在驱动力,人们认识到要更有效、更直接地反映图像视觉信息,就必须基于图像本身的属性。目前,基于内容的图像检索使用诸如颜色、纹理、形状、空间关系等图像的视觉内容来表示和索引图像。现在基于内容的图像检索已经融入了相关反馈技术,让用户参与到检索过程中,从而获得知觉、语义上更加有意义的检索结果。

11.3.3 图像特征的概念

图像的特征是人对图像视觉感受的量化描述。图像特征从各个方面描述了图像的内在语义,从而可以作为图像的抽象表示。基于特征的图像检索,或称为基于内容的图像检索,利用不同特征定义的相似度表示不同图像之间的相似程度,从而实现对图像的检索。可见,基于内容的图像检索首先要确定特征,以便让计算机自动地或半自动地从图像中提取这些特征。其次根据这些特征进行相似性度量,认为查询图像与目标图像特征值越接近,则两幅图像越相似。

一般把图像特征分为低层特征和高层语义特征。低层特征如上述的颜色、纹理、形状和空间关系等,它们简单、直观、有效,所以现有的大多数 CBIR 系统都采用了这些特征。高层语义特征描述了图像本身与其反映的客观世界之间的关系。由于图像内容的复杂性,人对图像内容的理解往往建立在经验的基础上,所以人对图像内容的理解很难仅靠统计特征来描述。语义特征或者由文本标注得到,或者经过基于视觉内容的复杂推理过程得到,因其复杂性而在现有的 CBIR 系统中很少得到应用。

11.3.4 基于内容的图像检索系统的框架

根据目前 CBIR 技术的发展和现有的 CBIR 系统的总结,可以建立基于内容的图像检索的基本框架,如图 11.8 所示。

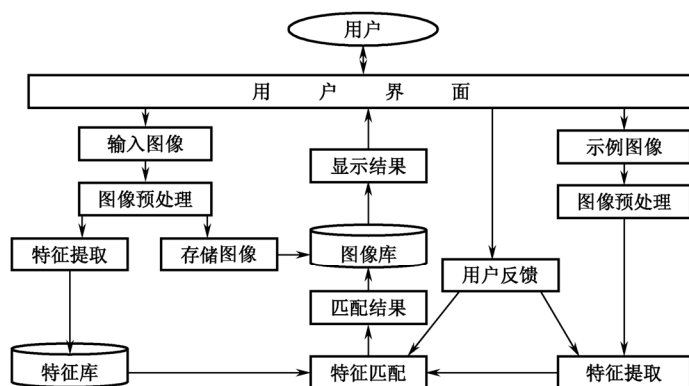


图 11.8 基于内容的图像检索系统

分析上图可以看出, CBIR 系统一般包括下面几个模块: 用户界面模块、图像输入模块、特征抽取模块、结果显示模块及用户反馈模块。各模块独立完成一定的功能, 用户界面模块提供系统与用户的接口; 图像输入模块在向系统输入图像的同时要对图像进行预处理, 并对图像进行特征抽取以建立特征库; 特征抽取模块完成对特征的提取, 这是决定 CBIR 系统性能优劣的关键之一; 结果显示模块把示例图像或草图的特征与图像特征库中的特征进行相似匹配, 并按相似度从大到小显示图像; 用户反馈模块完成用户与系统的交互, 用户可以人工标出相似图像作为正反馈、不相似图像作为负反馈, 系统根据用户的反馈对参数做相应的调整, 从而提高查询的正确率。

1. 图像的预处理

在我们把原始图像转换成数字图像过程中, 由于种种原因, 图像的画质可能会出现种种不尽如人意的退化。例如, 灰度偏移、几何变形及图像变得模糊不清等。所以, 抑制使图像退化的各种干扰信号, 增强图像的有用信号, 以及将观测到的不同图像在同一约束条件下进行校正处理, 就显得非常重要。常用的预处理方法有滤波、几何变换、直方图均衡、模糊图像的锐化、灰度值修正等。另外, 图像预处理还包括格式转换、尺度归一等。

2. 图像的特征提取

这是基于内容图像检索系统的关键技术之一, 特征提取的好坏直接关系到整个 CBIR 系统的性能, 对特征提取技术的基本要求是准确和快速, 特征选取时要考虑下列 4 个因素:

- (1) 图像的区分能力, 应能很好地区分视觉上差异较大的图像;
- (2) 一次查询中可能检索到的最大相关图像数, 即无关图像排除能力;

(3) 特征计算复杂度, 复杂度过高会影响系统反应时间;

(4) 特征的存储空间要求。

要建立 CBIR 系统, 首先要对图像内容进行分析, 抽取特征, 再用这些特征建立索引, 并进行检索。具体地说, 就是提取图像的视觉特征如颜色、纹理、形状等, 并用多维特征向量表示, 然后再把这些特征向量组成特征库以便快速地进行检索。比如使用颜色直方图来表示图像的特征, 假设量化等级为 256, 如果没有简化, 则特征向量高达 256 维。由于提取的图像特征向量的维数高达数十甚至上百维, 所以建立图像的特征库是必要的, 否则对于图像数据库特别是大型的图像数据库而言, 查询时间是令人难以忍受的。值得注意的是, 提取出的这些特征向量与图像之间往往不是一对一的映射关系, 而是一对多的映射关系, 所以图像特征提取性能很大程度上决定了图像数据库的性能。

3. 图像相似性度量

图像相似性度量就是对提取出的图像颜色、纹理、形状、空间关系等特征进行匹配。决定 CBIR 性能好坏的两个最重要的问题是: 第一, 如何利用图像的低级可视特征如颜色、纹理、形状等有效地表示图像的内容, 即上述的图像特征提取; 第二, 如何度量图像之间的相似性。在 CBIR 系统中, 图像之间的相似性通常是指它们的特征向量之间的相似性, 将图像的特征向量视为某个特征空间的点, 两点的接近程度用它们的距离表示, 距离越小则表示它们所代表的图像越相似。用户检索图像时, 首先向系统提供示例图像 (instance image) 或草图, 然后系统将示例图像或草图转换成其特征向量的内在表示形式, 接着系统计算用户所给图像或草图与图像库中图像特征向量的相似度或距离, 最后系统借助于索引机制实现检索。

与基于文本检索的精确匹配不同, CBIR 通过计算查询图像与库中图像的相似度来检索图像, 所以检索返回的结果不是单幅图像, 而是若干图像, 一般按照相似度从大到小显示图像以供用户选择。

4. 图像检索性能的评价

图像检索评价方法有很多, 较常用的是图像的查准率 (Precision) 和查全率 (Recall)。这种评价标准最初用于传统的信息检索系统中。查准率是指在一次查询过程中, 所查到的相关图像数目 T_1 同该次查到的所有图像数目 T_2 之比; 查全率指在一次查询过程中, 用户所查到的相关图像的数目 T_1 和数据库中目标图像相关的所有图像数目 T_3 之比。即

$$\text{Precision} = T_1/T_2 \quad (11.7)$$

$$\text{Recall} = T_1/T_3 \quad (11.8)$$

一般预先设定一个阈值, 当其相似度大于阈值时, 作为相似图像返回。在这个定义下, 查全率和查准率成反比关系: 提高阈值则返回的图像减少, 查全率降低而查准率提高; 反之亦然。所以要综合使用这两个指标来衡量某种检索算法。

使用查准率和查全率两个公式需要预先设定一个相似度阈值, 而这个阈值的确定是主观的, 特别是图像相似度的大小随着所取特征的不同而变化, 这就更为阈值的选择增加了难度。另外, 计算查准率时要考虑图像库中的所有相似图像, 如果图像库很大, 也会给我们统计相似图像带来一定的难度。

下面引入 MPEG-7 所推荐使用的另外一种评价检索性能的两个公式。这两个公式含义清晰、计算简单、使用方便。这里先设 P_N 为返回的前 N 个结果中的正确率, 又设 R 为某一具有特定语义含义的图像集合。现给出示例图像 $q_i, i = 1, \dots, M$ 。 M 是为测试系统检索性能而预先设定的示例图像个数, 存在 $q_i \in R$, 在对 q_i 的一次检索中系统返回的前 N 个结果为 $p_j, j = 1, \dots, N$, 那么正确率 $P_N(q_i)$ 定义为

$$P_N(q_i) = \sum_{j=1}^N \frac{\psi(p_j, R)}{N} \quad (11.9)$$

式中, $\psi(x, Y) = \begin{cases} 1, & x \in Y \\ 0, & x \notin Y \end{cases}$ 。那么对于所有示例图像 q_i , 就得到一个平均正确率, 定义为

$$p_N = \sum_{i=1}^M \frac{(P_N(q_i))}{M} \quad (11.10)$$

5. 用户相关反馈

虽然近十年来人们对基于图像内在属性的检索算法进行了各种研究, 但 CBIR 系统的检索效果仍差强人意。这主要是由于图像的低级视觉特征和图像的语义内容之间存在着很大的差别, 所以图像的低级视觉特征如颜色、纹理、形状等未必能正确在反映人对图像的理解。更何况, 不同的人对同一幅图像的语义解释各不相同, 即使同一人在不同时期对同一幅图像的理解也不一定相同。为了突破这个瓶颈, 近几年人们又在 CBIR 系统中引进了相关反馈技术, 把人们对图像的主观理解与基于特征的图像检索融合在一起, 让系统能自动地改良返回结果。相关反馈技术最初起源于文档检索系统, 但成果有限, 当这种技术引入 CBIR 系统后, 却得到了人们的极大关注, 原因就在于通常的 CBIR 系统检索效率较低, 而具有相关反馈技术的 CBIR 系统能显著地提高检索的精确度。

11.3.5 相似度测量公式

近年来, 在图像检索实践中, 人们基于对特征分布的经验估计研究了许多相似度测量公式。不同的相似度测量公式 (或称为距离测量公式), 对系统的图像检索结果影响很大。这里介绍几个在 CBIR 系统中广泛使用的相似测量公式。这里先假设 $D(I, J)$ 为查询图像 I 和库中图像 J 之间的距离, $f_i(I)$ 是表征图像 I 的特征向量的第 i 个分量处的特征值。

1. Minkowski-Form 距离公式

如果图像特征向量的每一维彼此独立且其重要性相同 (即取相同的权值), 那么使用 Minkowski 距离公式是较为合适的。公式定义如下

$$D(I, J) = \left(\sum_i |f_i(I) - f_i(J)|^p \right)^{1/p} \quad (11.11)$$

式中, p 可取 1、2 和 ∞ 。假设用 L_p 表示 Minkowski 距离, 则当 p 取 1、2 和 ∞ 时距离分别为 L_1 、 L_2 和 L_∞ , 其中 L_1 和 L_2 就是我们熟知的绝对距离 (街区距离) 和 Euclidean (欧氏) 距离。Minkowski 距离公式是图像检索中使用最为广泛的测度。

2. Quadratic Form (QF) 距离公式

Minkowski 距离把特征向量的所有分量视为完全独立且彼此不相关, 但没有考虑到这样的事实, 即特征向量的某几个分量在知觉上可能比其他分量更相似。为了解决这个问题, 引入了 QF 距离公式, 即

$$D(I, J) = \sqrt{(\mathbf{F}_I - \mathbf{F}_J)^T \mathbf{A} (\mathbf{F}_I - \mathbf{F}_J)} \quad (11.12)$$

式中, $\mathbf{A} = [a_{ij}]$ 是相似矩阵, 其中 a_{ij} 表示特征向量的分量 i 和 j 之间的相似度。 \mathbf{F}_I 和 \mathbf{F}_J 表示特征 $f_i(I)$ 和 $f_j(J)$ 的向量。由于 QF 距离公式考虑了颜色之间的相似度, 所以它计算出的结果比欧氏公式和直方图相交公式更好地反映了人的知觉特点。

3. Mahalanobis (马氏) 距离公式

马氏距离公式适用于图像特征向量的各维彼此不独立, 且其重要性各有不同的情况。其定义如下

$$D(I, J) = \sqrt{(F_I - F_J)^T C^{-1} (F_I - F_J)} \quad (11.13)$$

式中, C 是特征向量的协方差矩阵。

4. Kullback-Leibler (KL) Divergence 和 Jeffrey-Divergence (JD) 距离公式

KL 距离公式是用一个图像的特征量作为码本来反映另一个图像的特征分布情况, 两图像 I 和 J 的 KL 距离公式定义如下

$$D(I, J) = \sum_i f_i(I) \log \frac{f_i(I)}{f_i(J)} \quad (11.14)$$

JD 距离公式定义如下

$$D(I, J) = \sum_i (f_i(I) \log \frac{f_i(I)}{\hat{f}_i} + f_i(J) \log \frac{f_i(J)}{\hat{f}_i}) \quad (11.15)$$

式中, $\hat{f}_i = [f_i(I) + f_i(J)]/2$ 。

KL 和 JD 距离公式比较适合于纹理特征的相似性测量。

11.3.6 基于内容的图像检索系统简介

自 20 世纪 90 年代以来, 基于内容的图像检索研究成为一个关注的热点, 因而出现了许多基于内容的图像检索系统, 较著名的有 IBM 公司的 QBIC (Query By Image Content) 系统、VIRAGE 系统和 VisualSeek/WebSeek 系统等。这里简要介绍国内外的几个典型系统。

1. QBIC 系统

QBIC 是 IBM Almaden 研究中心于 20 世纪 90 年代开发的, 它的成功开发对后来的图像检索技术发展产生了深远影响。

QBIC 提供了基于颜色特征、纹理特征和形状特征的检索功能。基于颜色特征的查询主要采用颜色百分比查询和颜色直方图查询两种方法; 基于纹理特征的查询则综合使用了图像中线条的粗糙性、对比性和方向性等三种特征; 基于形状特征的查询采用了形状的面积、离心率、主轴方向及一组变换无关量等描述方法。

虽然 QBIC 系统可以使用多种特征进行检索, 但这些特征间是孤立而没有联系的, 同时用户使用该系统需要具有图像特征的一些基本知识, 这就使得该系统的易用性大打折扣。

2. VIRAGE 系统

VIRAGE 提供了 4 种可视特征检索, 即颜色、颜色布局、纹理和形状。每种特征被赋予 0~10 的权值。用户需要设定一个或多个属性权值来优化检索, 要达到最佳平衡度, 需要反复实验。所以同 QBIC 系统一样, 存在着特征选择盲目、用户不易使用等缺点。

3. VisualSeek/WebSeek 系统

VisualSeek 和 WebSeek 是美国哥伦比亚大学开发的图像检索系统, 前者是视觉特性搜索引擎, 后者是面向 WWW 的文本/图像搜索引擎。

VisualSeek 实现了基于颜色、形状、空间关系以及文字关键字的查询。WebSeek 侧重于 Web 应用，主要有三个模块：图像、视频收集模块；分类和检索模块；浏览和检索模块。

近几年来，国内在这方面也取得了一些进展，如章毓晋等人研制成功了一种基于特征的图像查询和检索系统。该系统不仅提供了多种包括颜色、纹理、形状等单一特征的查询和检索手段，而且还可综合利用各种不同特征进行检索操作。系统采用的特征包括基于累积直方图和局部累积直方图，这是颜色特征；基于灰度共生矩阵的 4 种纹理量；基于不变矩的形状特征；等等。系统能自动提取特征向量，能同时显示查询图、特征图和检索结果。

11.3.7 基于内容的图像检索技术的研究热点

基于内容的视觉信息检索（包括静态图像、视频等）从 20 世纪 90 年代开始，已经过十多年的发展，取得了巨大的成就，也在各个领域得到广泛应用。但作为一个新兴的、具有挑战性的课题，这方面的工作毕竟还处于起步阶段，仍然具有巨大的研究和应用空间。同时这方面的工作在经过十多年发展后也暴露了许多问题，所以很多人都在关注这个领域今后的发展。目前人们试图从多方面来提高检索性能，研究得较多的有综合特征检索技术、相关反馈检索技术以及基于语义的检索技术等。

考虑到颜色、纹理、形状和空间关系等特征各有特点，它们各自反映图像某一方面的内容，如果进行综合利用则有可能提高检索性能。相关反馈技术实现了人机交互，用户根据先前的检索结果借助权重调整已有的查询要求，以便给检索系统提供更多、更直接的信息，从而使系统能更好地满足用户的需求。上述两种方法都是基于图像的低级特征，从人的认知角度来看，人对图像的理解主要是在语义层次上的，所以如果能将图像的语义特征结合到检索系统中，将会极大地提高检索性能。

将 CBIR 技术与网络技术相结合对视觉信息检索技术又提出了新的挑战。大量图像数据库与网络相连接，而且网络本身也有海量的图片，使图像搜索引擎成为当今的研究热点之一。

CBIR 技术的应用范围越来越广，涉及很多领域。指纹、人脸的识别可用于身份的认证；商标的查询可用于查找商标库中是否存在相同或相似的商标，便于商标的注册；特别是在医学领域，医生通过医学图片查找过去同类病情的治疗方案，以解决某些疑难杂症等。随着社会的进步和技术的发展，新的应用领域也在不断拓展，这给 CBIR 技术提出了新的要求。

11.3.8 一种基于颜色和纹理特征的彩色图像检索算法

1. 概述

近年来，对于图像内容描述的算法层出不穷，也相继推出了许多基于内容的图像检索系统。由于颜色是人识别图像的主要感知特征之一，因而基于颜色的查询是基于内容的图像检索中最基本的算法。常用的颜色索引技术包括颜色直方图（color histogram）、颜色矩（color moment）等。目前，颜色直方图仍然是基于内容的图像检索中最流行的颜色表示算法之一。除了图像的颜色特征外，纹理信息是另一种重要的图像低层特征。纹理通常定义为图像的某种局部性质，或是对局部区域中像素之间关系的一种度量，可用来对图像中的空间信息进行一定程度的定量描述。常见的如基于空间灰度的共生矩阵、高斯马尔可夫随机场模型等。

本节提供了一种有效的基于图像颜色和纹理特征的图像表示算法及相应的检索技术。小波变换能够将图像分解为不同尺度和方向的频带，其中，低频子带集中了图像的主要信息，有研究表明，一幅图像 95% 以上的能量集中在低频。对于颜色特征，在对 HSV 颜色模型进行特殊处理的基础上，采用 HSV 颜色空间的直方图；对于纹理特征，由于纹理特征大都集中在图像的中、高频部分，可以用小波

变换后的多尺度信息来描述。因此,利用图像小波分解后不同频带的颜色和纹理特征对花卉图像库进行检索,可以取得很好的检索效果。

2. 颜色特征的提取

(1) 颜色空间

HSV 颜色空间直接对应于人眼色彩视觉特征的三要素,并且从 RGB 到 HSV 的转换是一个简单且快速的非线性变换。

为了使 HSV 颜色模型能更好地符合人眼的视觉特性,首先将图像中每一像素的 R, G, B 值转换为 H, S, V 值 ($H \in [0, 2\pi], S \in [0, 1], V \in [0, 1]$), 然后根据它的特性做如下特殊处理:

- ① 黑色区域: 所有 $V < 15\%$ 的颜色均归入黑色, 令 $H = 0, S = 0, V = 0$;
- ② 白色区域: 所有 $S < 10\%$ 且 $V > 80\%$ 的颜色归入白色, 令 $H = 0, S = 0, V = 1$;
- ③ 彩色区域: 位于黑色区域和白色区域以外的颜色, 其 H, S, V 值保持不变。

经过 HSV 颜色变换, 图像中每个像素的颜色用 H, S, V 值表示。由于将与黑色、白色相近的颜色分别作为同一种颜色对待, 进一步提高了颜色模型的准确性。

(2) 颜色直方图

用 HSV 空间的色直方图来描述图像的全局颜色特征。一幅图像的颜色一般很多, 因此直方向量的维数会非常多。因此, 这里采用 HSV 颜色模型来对 H, S, V 这 3 个分量按人的颜色感知进行非等间隔的量化, 在对颜色模型进行大量分析后, 可把色调 H 分成 16 份, 把饱和度 S 和亮度 V 分别分成 4 份, 把 3 个颜色分量合成为一维特征向量, 即

$$L = 16H + 4S + V \quad (11.16)$$

式中, 对色调 H 、饱和度 S 、亮度 V 分别取权重为 16、4、1, 这就大大减轻了图像亮度 V 对检索结果的影响, 同时也减少了饱和度 S 对检索结果的影响, 但对颜色分布不同的图像却能很好地检索出来, 故能充分利用图像的颜色信息特征来满足人们对图像查询的要求。根据式 (11.16) 可知, L 的取值范围是 0~255。

1994 年, W. Sweldens 等提出了全新的提升算法来构造整数小波。它不但保持了传统小波的主要特性, 而且适用于更为一般的不规则样本集合, 得到了广泛的应用。在提取颜色特征之前, 可以先对图像做整数小波变换, 然后从低频按照上述算法提取颜色直方图。由于每做一次小波变换, 就得到一个只有原图 1/4 大小的低频部分, 这就大大降低了计算复杂度。

3. 纹理特征的提取

由于全局颜色直方图只记录了全局的颜色统计信息, 丢失了颜色的空间分布信息, 可以通过纹理特征来弥补这一不足。众所周知, 图像的细节区域是图像的重要特征, 在图像检索中具有非常重要的作用。

对原图像进行二维整数小波分解后, 可以得到 4 幅子图像。其中, LL 代表低频部分, HL、LH、HH 代表 3 个高频部分。图像的纹理特征反映在高频部分, 因此可以采用图像小波分解后高频部分的信息来描述纹理。

研究表明, 2/10 整数小波变换能很好地提取图像的高频部分。由于图像的小波系数并不能直接用于描述图像的特征, 因此需要对其进行必要的统计。这里采用三级小波分解后不同尺度和不同方向高频部分的方差。

定义: 设 $f(x, y)$ 表示原始图像, $D_{2^{-i}}^1 f(x, y)$, $D_{2^{-i}}^2 f(x, y)$, $D_{2^{-i}}^3 f(x, y)$ ($i = 1, 2, \dots, N$) 表示图像第 i 次分解后的 3 个高频小波系数矩阵, 纹理特征向量定义为

$$V_{\text{texture}} = \left[\sum_{j=1}^3 R(D_{2^{-1}}^j f), \sum_{j=1}^3 R(D_{2^{-2}}^j f), \dots, \sum_{j=1}^3 R(D_{2^{-N}}^j f) \right] \quad (11.17)$$

式中, $R(\cdot)$ 代表方差。纹理特征即第 i 次小波分解高频部分 3 个系数矩阵的方差之和。

4. 距离度量与相似检索

(1) 子特征的距离度量

对于颜色特征, 采用直方图相交运算进行度量:

$$d(P, Q) = \frac{\sum_{l=1}^K \min\{P_l, Q_l\}}{\sum_{l=1}^K Q_l} \quad (11.18)$$

式中, K 为特征的维数, P 为查询例图, Q 为图像库中的一幅图像, $d(P, Q) \in (0, 1)$, 为两幅图像的匹配值, 它的值越接近于 1, 则两幅图像越相似。

对于纹理特征, 可采用比值相似度的算法进行度量。第 k 个分量的相似度定义为

$$\alpha_k = \begin{cases} 1, & V_k^P = V_k^Q \\ \frac{\min(V_k^P, V_k^Q)}{\max(V_k^P, V_k^Q)}, & \text{其他} \end{cases} \quad (11.19)$$

两幅图像之间的相似度定义为

$$\alpha = \sum_{k=1}^M w_i \alpha_i \quad (11.20)$$

式中, M 为特征的维数, $\sum_{k=1}^M w_i = 1$, $w_i = 1/M$, 即 α 是各分量相似度的加权平均值。

(2) 综合特征的距离度量

当采用综合特征进行检索时, 由于提取的图像特征不同, 相应的距离度量也不同。因此, 为了使不同的子特征具有可比性, 需要对子特征的相似距离做归一化处理。这里采用通常的高斯模型, 则整体相似度为归一化距离的加权相似度之和

$$S = w_c S_c + w_t S_t \quad (11.21)$$

式中, S 为整体相似度, w_c 和 w_t 分别代表颜色相似度和纹理相似度的权重, S_c 和 S_t 各代表两幅图像的颜色特征和纹理特征的归一化相似度。

5. 实验结果

对于颜色特征、纹理特征以及综合特征 (颜色和纹理特征), 在花卉图像数据库中进行了检索。实验评价标准采用查全率和查准率。

图 11.9 显示了图像库中的部分测试图像。图 11.10 是对其中一幅花朵图像采用 3 种不同的特征进行检索所得的结果。由此可见, 综合特征的查询效果明显优于单个特征的查询。此外, 该检索算法还具有平移、旋转、尺度以及镜面变换的近似不变性。

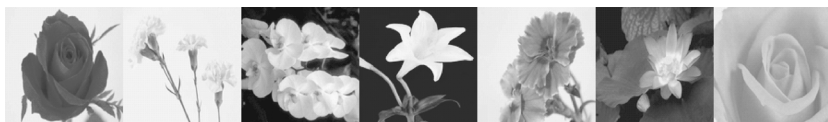


图 11.9 花卉图库中的部分测试图像

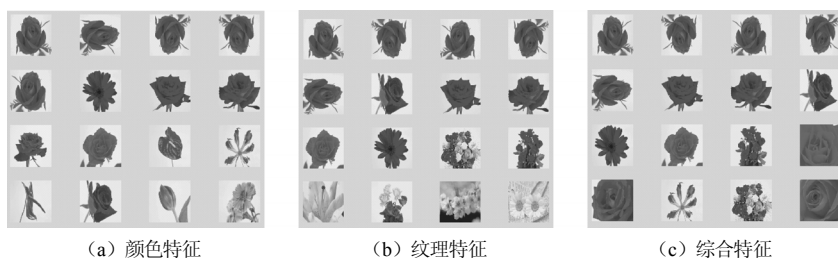


图 11.10 不同特征的检索结果

对于权重 w_c 和 w_t 的取值, 没有具体的规定, 可以根据用户的需要而定。简单来说, 如果查询例图为纹理图像, 则 $w_t > w_c$; 而当例图为非纹理图像时, 则 $w_t < w_c$, 且满足 $w_c + w_t = 1$ 。

通过多次实验结果的统计, 可以得出, 采用综合颜色和纹理特征对图像数据库检索时, 当查全率为 0.9 时, 查准率在 0.8 以上; 而当检索的查全率为 0.6 时, 查准率在 0.9 以上。实验表明, 上述的算法快速、有效地实现了彩色图像的相似检索。

6. 小结

基于颜色和纹理特征的图像检索算法, 采用了改进后的 HSV 颜色直方图和整数小波变换后的高频部分的方差统计量。HSV 颜色直方图无法描述图像的空间分布信息, 而整数小波的多尺度信息正好弥补了这一缺陷。实验结果表明, 采用两种特征相结合的算法进一步提高了图像检索的性能。

11.4 数字化医院中的图像存档与通信系统 (PACS)

11.4.1 概述

在没有统一的工业标准之前, 各厂商生产的医疗设备, 其产生的图像数据格式、数据传输方式各不相同, 这样就大大阻碍了图像数据及其相关信息在相关设备之间的交换, 给应用带来了极大的不便。基于上述原因, 美国放射学会 (ACR) 和美国电气制造商协会 (NEMA) 在参考了其他国际标准的基础上, 联合推出了医学数字图像存储与通信标准, 即 DICOM 标准。在 1993 年正式推出 DICOM 3.0 标准之前, ACR 和 NEMA 首先于 1985 年推出了 ACR-NEMA 1.0 标准, 然后在 1988 年推出了 ACR-NEMA 2.0 标准。这样, 各厂商只要遵从 DICOM 标准, 产生的医学影像数据及其相关信息就能方便地进行交换。对于原有的不符合标准的设备, 医院可以开发专门的接口, 利用 DICOM 标准计算机联网技术, 就可以与另一台支持 DICOM 标准的设备直接相连。同时, DICOM 标准的推出和实现也推动了远程放射会诊、图像档案管理和通信系统 (PACS) 的研究与发展, 并且由于 DICOM 的开放性和互连性, 使得与其他医学应用系统 (HIS、RIS 等) 的集成成为可能。

随着社会的发展和技术的进步, 人们对身体健康关心的程度越来越高。医学影像已经不再是仅供医生参考的信息, 而是已成为诊断疾病的重要依据; 同时医院的影像科也已从辅助科室成为救死扶伤的生力军。电子病历数字化是进入医院的根本标志, 而建立电子病历当数图像的存储、检索和通信最

为困难。1992年,第一个集文字和图像于一体的电子病历系统——图像存档和通信系统(Picture Archiving and Communication System, PACS)在美国诞生。PACS是图像信号处理技术和计算机通信技术密切结合的典型产物,是具有包括图像获取、存档、检索、传送、显示、管理等功能的完整的网络系统。

网络条件下的图像压缩编码技术是PACS的重点,也是21世纪图像通信和远程计算机网络的关键技术。PACS的目标与办公自动化中的“无纸办公室”目标对应,旨在实现“无胶片医院”。用多媒体计算机实现医学影像的数字化存档和管理,使医护人员可以克服地域和时间上的限制,根据访问权限通过Intranet查询和利用本院的影像资料,而在本市不同医院或异地医院可以通过城域网、Internet进行远程访问和专家会诊。事实上,在中小城市,价格昂贵的数字化医疗设备,靠一家医院是很难承受的,PACS的建立为在类似异地诊断、本地护理的医疗模式提供了技术设施。

PACS对提高医疗质量,加强医疗、科研、教学、管理,促进医院信息化建设有重要意义,也是发展地方医疗事业信息化和数字化城市的重要组成部分。总之,PACS在城市医院中的建立对于改进病人的护理质量、节省医务人员的服务时间、满足医疗个性化等多方面具有显著的社会效益和潜在的经济效益。

11.4.2 国内外发展现状

PACS按照规模和功能分,有小型PACS、数字PACS和全规模PACS。它们分别以服务于单一影像部门、服务于含X线影像及其所有影像设备和服务于全放射科或医学影像学科范围甚至包括临床影像浏览、会诊系统和远程放射学服务为特征。目前国外已经处于“第二代PACS(Hi-PACS, Hospital integrated PACS)”阶段。Hi-PACS的基本含义包括模块化结构、开放式架构、DICOM标准、整合HIS-RIS(医院信息系统-放射学信息系统)等特征的全规模PACS。

典型的工作有日本HOKKAIDO大学的HU-PACS系统。该系统1989年就开始为临床服务。开始的服务仅限于放射科和门诊部,实际上相当于一个RIS。自1994年起开始第二阶段的建设,通过安装CT和MRI扩大服务到全部病房及管理图像数据。由于包括图像在内的医院信息迅速增长,建立基于DICOM 3.0标准的开放式系统刻不容缓。第三阶段的目标是用ATM技术以星形网络拓扑结构将网络服务器数据库在几个计算机中分散负载,完成全院范围的服务和HIS-RIS-PACS的集成。图11.11是HU-PACS的体系结构图。

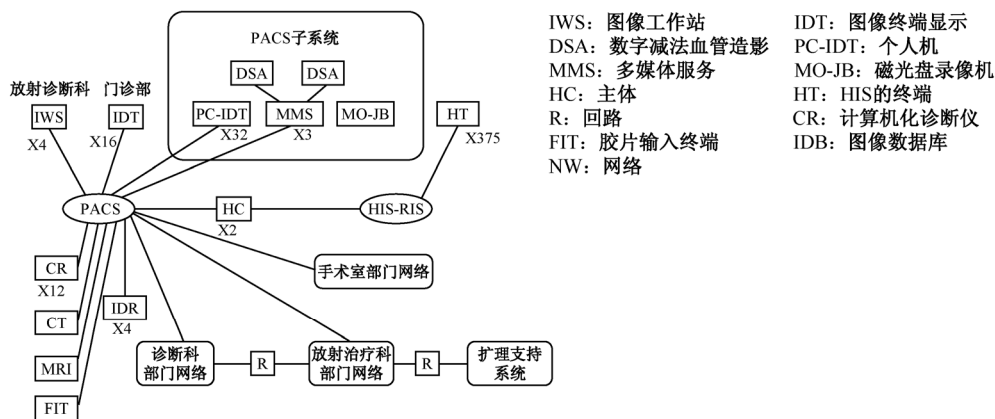


图 11.11 HU-PACS 的体系结构

国内在PACS的实践上也已经取得了许多应用成果。遵循高起点、标准化设计、根据投资状况及对技术的驾驭程度阶段性实施,是符合国情的PACS发展原则。值得推荐的有标准框架设计模式(见

图 11.12) 和单一平台设计模式两种。前者代表了 PACS 的发展方向, 尽管代价高, 但便于实现分布式管理, 性能好; 后者以影像服务器为中心, 容易构建集中式管理模式, 设备利用率高, 成本较低, 但可扩展性差, 对未来的系统升级不利。

今后 PACS 的发展趋势是数字化 PACS (digital PACS)、小型 PACS (mini PACS) 向全规模 PACS (full-service PACS) 过渡, 实现 HIS-RIS-PACS 三大医院信息系统的集成和通信, 可以为远程交互影像会诊服务体系的设计和规划建立基础。由于医疗事业涉及千家万户, 数字化医院自然得到医务人员、政府和市民等各阶层有识之士的关注, PACS 已成为数字化医院的热点研究课题。

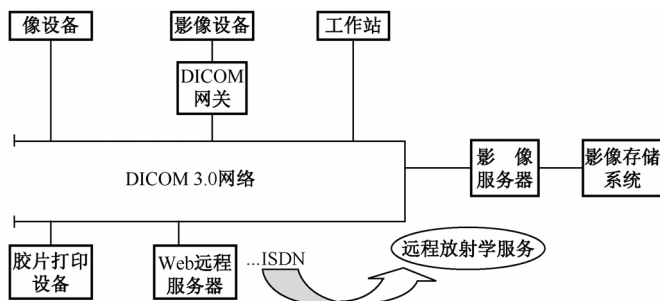


图 11.12 PACS 标准框架设计模式

11.4.3 主要解决的问题和技术要点

PACS 的建立投资大, 周期长, 学科交叉性强, 影响面广, 所以科学的方案论证尤为重要。方案论证主要解决以下问题。

(1) 投资规模问题。从国外的经验来看, PACS 的资金投入从数万美元到上百万美元不等。应该根据国家和地方的实际情况建立 PACS 的体系结构, 分阶段实现数字 PACS、小型 PACS、全规模 PACS, 最终实现 HIS-RIS-PACS 三大医院信息系统的集成和通信, 为电子病历、远程放射学和远程交互影像会诊服务体系服务。

(2) 基于 DICOM 的 PACS 与现存信息系统的兼容性问题。早期的 HIS-RIS 没有遵循医疗保健的 HL-7 等国际标准, 使系统今后的扩充带来了很大的困难。新建的系统一定要遵循 DICOM、MPEG-7 等医学图像、多媒体等国际标准。DICOM 标准遵从性 (compliance) 是当前各国医学图像网络的基本要求。由于 DICOM 标准具有开发性, 则对于有不同影像设备供应商提供的设备组成的 PACS 必须确定完整的 DICOM 标准遵从性、DICOM 各子系统的互连 (interconnectivity) 和互操作性 (interoperability)。图 11.13 所示为 DICOM 3.0 标准的体系结构。DICOM 3.0 标准采用了面向对象技术, 为开发 PACS 软件提供了一个重要标准。该标准规定了医学图像数据及对它们的操作, 规定了数据结构的语义和数据字典, 也规定了数据存储和交换的格式。用 DICOM 3.0 格式的数字信号, 可以实现多个文件传输的同时性。采用 OSI 技术和网络协议应用层的约定, 可以实现医学图像无视觉损失的存储和传输。

(3) 研究医学图像的数字化、影像设备的数字接口问题。各种医学影像设备 (CT、MRI、PET、SPECT、B 超、心电图、胃镜、DSA 等) 都必须进行数字化处理, 为了保证图像的质量、提高诊断的准确度、避免治疗事故的法律纠纷, 可采用 DICOM 转换器等网关设备获得各种数字格式的高保真医学影像。对已有影像资料可以采用胶片扫描仪等设备进行数字化处理。注意 DICOM 与 HL-7 国际标准的接口、DICOM 和 HL-7 与非标准系统的接口的连接。

(4) 压缩存储和传输, 提高医学图像的复用率。采用小波变换、分形等模型编码方法可以进行图像数据的有效压缩, 且压缩比可以根据传输介质的不同进行分层选择。新的编码方法充分利用景物中

的内容，突破了传统信源编码理论的框架。相对于一般的数字化过程来说，医学影像的保真度十分重要，满足临床诊断要求的影像压缩是图像压缩编码的新领域。现在已有支持 DICOM 3.0 格式的 2:1 无损、50:1 有损实时图像压缩/解压缩卡。高品质小波压缩卡可以实现 200:1 的压缩。人体器官的组织结构是较强的，采用分形来重组图像实现压缩将获得极大的压缩比，特别适合在网络中传输。如果分形与小波变换等其他压缩方法结合，将扬长避短，使压缩性能获得进一步提高。PIFS 码利用图像区域间的相似性，为 IFS 码的自动生成、半自动生成、失真度准则及实时编码等提供了新的途径。一般复杂度的图像以迭代和随机规则为特征函数，则用几百个 IFS 码就可以完全描述，但目前其编码过程仍依赖于人机交互方式。可以考虑用 DSP 模拟高效算法并固化到 ASIC（专用集成电路）中实现实时化。

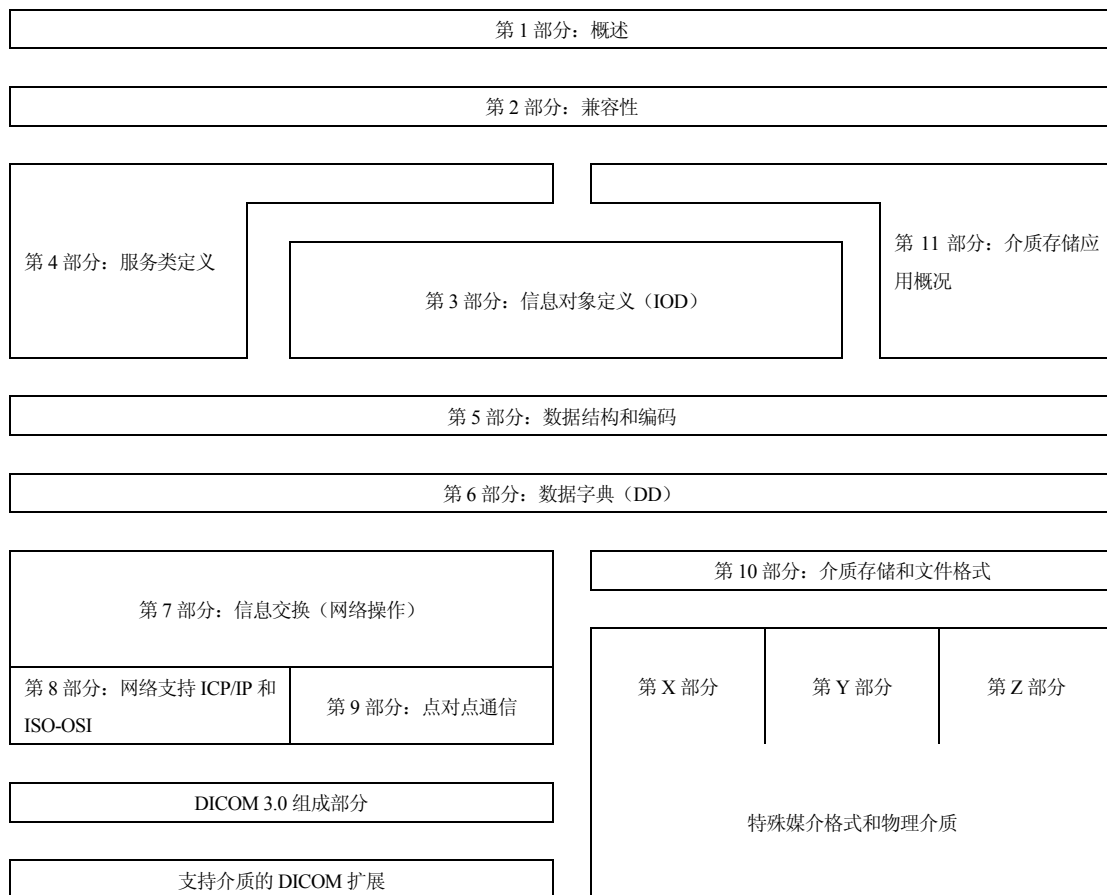


图 11.13 DICOM 3.0 标准的体系结构

(5) 图像数据库的建立和图片检索问题。医学图像服务器主要用来存储医学影像信息，响应用户索取图像的要求，管理 Intranet 的工作，它应该支持分布式和多媒体的要求。符合 MPEG-7 的 CMIR 技术可以用于 PACS 系统。

(6) 图像终端显示和浏览。在高档个人机上，以友好的人机交互方式开发影像浏览软件。软件能够快速显示、计算、分析图像，对图像有冻结、放大、测量、计算灰度（灰阶）面积和直径等功能。只有方便的操作方式才能使 PACS 的优势通过医务人员的使用得到充分体现。

在研究和建立 PACS 中有几个值得注意的原则如下：

(1) 在资金或硬件条件暂不具备的地区，可以设立充分利用现有数据采集等过程中的医疗设备并

进行 DICOM 化。密切注意数字化设备和多媒体、网络技术的发展,分阶段实施。HU-PACS 分三期的分步实施策略就是一个范例。

(2) 按系统工程和面向对象软件工程的原则开展研究和开发工作。借鉴国内外已有的理论和方法,避免重复研究工作,在短期内达到较高的研究水平。对本院和其他医院现有的 PACS、HIS 或 RIS 进行充分调研和深入研究,以便确定最佳方案。国内不少医院在建立信息系统中带来了效益,但也有些系统运行不理想或产生的经济和社会效益甚少,重复投资现象时有发生。究其原因,主要是在决策规划时没有遵循统筹兼顾的原则。在考虑方案时,可以考虑“院(医院)、校(大学)、政(政府)”的联合,使项目的理论和技术基础、仪器设备、资金等充分得到保障。充分听取医师、计算机技术人员、医疗器械工程师、管理人员等各类人的意见,使方案建立在科学性、可行性的基础上。

(3) 探讨多种 PACS 方案,根据资金和现有设备的制约情况提供可供实施、优化的数字 PACS、小型 PACS 和全规模 PACS,分析各自的性能,做到求新务实。在部分实施时特别要考虑到方案的前瞻性,全盘考虑现状和未来的发展,留有足够的发展空间。

(4) 静态的(压缩存储)研究与动态的(如心脏等器官的实时视频扫描图像的网上传输)研究相结合,以 DICOM 3.0 标准统一整个系统。区分服务对象、不同图像质量、不同图像特征等问题来考虑供教学、科研、临床等方面的不同应用,分类存储和传输。

(5) 将服务器数据库的建设与工作站浏览器的开发相结合。局域网是 PACS 的硬件基础,建立基于 TCP/IP 网络通信协议和 WWW 规范的 Intranet 是符合潮流的。图像数据库在电子病历的地位举足轻重,是医院信息系统中的核心部分。工作站浏览器是面向用户的最终软件,其质量将直接影响到用户对系统的评价,同样不可轻视。

11.4.4 DICOM 图像格式

DICOM (Digital Imaging and Communications in Medicine) 指医学数字成像和通信系统。DICOM 标准是由美国放射学会 (ACR) 和国家电气制造商协会 (NEMA) 共同制定的。从 1993 年推出 DICOM 3.0 后,就不断对它进行扩充和完善。DICOM 涵盖了数字图像信息构成和通信两个领域,内容极其庞大。现在最新的版本共分为 16 章。第 1 章和第 2 章是它的导论和一致性声明部分,它为我们描述了本标准所涉及的范围,标准的目的、宗旨、意义,以及参考了哪些其他国际标准。DICOM 标准允许工程实现者部分或全部地遵从,所以在产品说明中必须声明遵从的程度。标准的后几章讲述的内容包括端到端通信、数据交换中的存储和介质格式、端到端支持的打印管理、灰度标准显示函数、安全规范等。

用户可以根据医学工作的需要,制定符合 DICOM 标准的图像文件格式。医学图像文件结构部分主要参考标准的第 3 章:信息对象定义 (Information Object Definitions)、第 4 章:服务类规范 (Service Class Specifications)、第 5 章:数据结构和编码规定 (Data Structure and Encoding)、第 6 章:数据字典 (Data Dictionary) 和第 10 章:数据交换中的介质存储方式和文件格式 (Media Storage and File Format for Data Interchange)。

11.4.5 DICOM 3.0 标准及其面向对象的实现

下面主要分析最新的 DICOM 标准在体现面向对象方面的主要思想,在结合面向对象分析、设计和编程技术的基础上,就该标准的实现提供图像文件读取、显示和格式的转换等解决方案。

1. DICOM 信息对象定义和服务类

DICOM 标准融入了面向对象的思想。标准的信息对象定义 (Information Object Definition, IOD) 就是一个面向对象的抽象数据模型,它反映的是某类具有相同性质或属性的客观世界对象的信息。IOD

代表的不是客观世界对象的一个实例,而是具有相同属性的对象的类。可见,IOD 与面向对象编程中“类”的概念非常相似。DICOM 标准定义了信息模型来反映 IOD 之间的关系,这种信息模型是在表示 DICOM 标准适用范围之内所有现实世界对象之间关系的应用模型基础上抽象而来的。

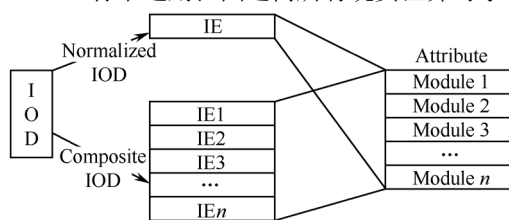


图 11.14 IOD 结构示意图

IOD 具有明显的层次结构 (见图 11.14), 这种结构为面向对象设计提供了可能。每个信息对象定义 (IOD) 由一个或多个信息实体 (Information Entity, IE) 组成, 信息实体对应着上述应用模型中的现实医疗实体 (如患者、图像等) 的一个数据抽象。每个信息实体又由若干属性 (Attribute) 组成, 属性是对医疗实体性质的抽象。如患者有姓名、性别、编号等属性

性, 而图像有成像设备、成像日期、图像的行数、列数等属性。DICOM 标准又把一个实体中相关的属性组合在一起形成可复用的模块 (Module)。为了满足标准在未来的发展需要和与以前版本的兼容, 标准把 IOD 分为两类: 规范 IOD (Normalized IOD) 和复合 IOD (Composite IOD)。规范 IOD 如患者 IOD、检查 IOD、打印队列 IOD 等仅包含一个信息实体, 因为它们所包含的属性都是实体本身所固有的。比如患者 IOD 中可以包含患者姓名、患者生日等患者所固有的属性, 而像成像设备类型、检查日期等属性不是患者所固有的, 因此不能包含在患者 IOD 之中。与此相比较, 复合 IOD 如图像 IOD 包含了一个以上的相关信息实体, 复合 IOD 大多为医学图像。图像 IOD 除了包含图像实体本身的属性外, 还包括像患者 (Patient)、检查 (Study)、设备等实体的属性, 后面的几个实体虽然不是图像 IOD 本身所固有的, 但与图像关系密切, 不可缺少。

DICOM 标准不仅用 IOD 对对象属性进行了封装, 而且也对处理这些对象的方法进行了封装。DICOM 把特定的服务与相关的 IOD 定义为一个服务对象对 (Service Object Pair, SOP), 这组特定的服务在 DICOM 中称为 DICOM 消息服务元素 (DICOM Message Service Element-service, DIMSE)。一个服务类由若干相关的对象对类组成。与 IOD 相似, 服务对象对也分为规范 IOD 和复合 IOD 两类。前者是由规范 IOD 加一组服务构成的, 后者则是由复合 IOD 加一组服务构成的。DICOM 标准共定义了八种服务类, 其中四种是复合服务类, 都是面向图像的; 另四种是规范服务类。复合服务类主要用于图像的存储、查询/检索、检查内容通知等服务, 以及对两个类之间是否有联系进行验证。规范服务类包括患者管理服务类、检查管理服务类、结果管理服务类以及打印管理服务类。

2. DICOM 标准数据结构、编码和文件格式

信息对象定义和服务类是对客观世界信息对象的抽象, 当具体到某一实例时就涉及编码问题。标准的第 5 章详细说明了 DICOM 数据结构和编码问题。根据面向对象的思想方法, 客观对象是由若干属性及操纵这些属性的方法构成的。DICOM 标准用数据集来表示一个信息对象实例, 而一个数据集又由若干数据元素构成, 性质相同的数据元素被归为一组, 每个数据元素与信息对象的一个属性一一对应 (见图 11.15)。数据集、数据元素组和数据元素之间的关系对应着信息实体、属性模块和属性。

由图 11.15 可以看到, 一个数据元素可以包括 4 个域: 标签 (Tag)、值表示 (Value Representation, VR), 值长度 (Value Length) 和值域 (Value Field)。其中 VR 是一个可选域, 用它来说明值的类型, 它的出现与否与传输语法有关。当 VR 出现时, 称之为显式 VR, 不出现时称为隐式 VR。标准详尽定义了 26 种 VR, 每种类型适用的字符集和相应于某一类型数据元素值的长度或最大长度。在数据集中一个数据元素由标签唯一确定, 它们之间是一一对应的关系。使用者定义的数据元素必须满足第 6 章数据字典的规定。例如, (0010, 0010) 这个数据元素存放患者姓名, 而 (7EF0, 0010) 中则存放图像像素。标签由两部分构成: 组号 (前者) 和元素号 (后者)。标准把具有相关性质的数据元素组成一组,

它们具有共同的组号,元素与元素之间的差别是由不同的元素号来标识的。如组号为 0010 的组中存放了有关患者的信息,其中元素号为 0000 的记录该组的长度,元素号为 0010 的记录患者的姓名,而元素号为 0020 的则记录了患者编号,等等。从这里也可以看出 DICOM 的层次性:一个数据集由若干元素组成,每一组又是由若干性质相关的数据元素构成的。数据集中各数据元素是按组号和元素号的升序排列的,每个数据元素在一个数据集中只能出现一次(嵌套除外)。数据元素可分为两大类,一类是表示对象属性的数据元素,另一类则是操作这些属性的命令元素。

数据元素中值长度域以字节为单位表明此元素中值域的长度,其值的长度必须为偶数。值域中包含了属性的具体值,其长度是可变的,但最大不能超过值长度域所标定的长度,且值的长度必须为偶数字节。如不是偶数,则根据规则补空格键(20H)或 NULL(00H)。值域也可以包括多个值,这就是所谓的值多样性问题(Value Multiplicity),值与值之间用反斜杠“\”隔开,其总长度(包括“\”)必须为偶数。

标准以文件形式封装一个服务对象对实例,若干文件组成文件集(见图 11.16)。DICOM 文件是由文件元信息和 DICOM 数据集构成的。前面我们已经详细介绍了 DICOM 数据集,这里简述一个元信息的构成。元信息由文件头和组号为 0002 的元元素构成。文件头包括 128 字节的文件前文 4 字节的前缀,这个前缀的内容只能是“DICM”,它用来表示此文件是 DICOM 格式文件。应用程序可以使用 128 字节的文件前文,也可以不使用,如不使用则填充“00H”。文件元元素使用显式 VR 表示,包括文件元信息版本、传输语法等内容。

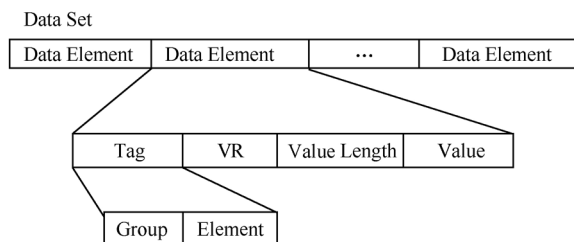


图 11.15 数据元素结构

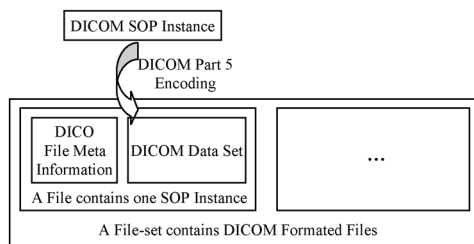


图 11.16 DICOM 文件集和文件格式

3. DICOM 文件的显示和格式转换

(1) DICOM 文件的显示

由于 Windows 系统仅支持位图格式图像的显示,因此其他格式的图像只有转化为位图格式才能在计算机中显示。人们熟知的 DIB 位图由信息头、颜色表和图像数据三部分组成,如果再加上一个文件头,那就是 BMP 格式文件。DICOM 文件中所含的信息远多于 DIB,只要把 DICOM 相关数据元素转换成 DIB 数据结构相应的部分就可以了。例如,对 DICOM 文件数据元素(0028, 0010)、(0028, 0011)、(0028, 0100)分别表示图像的行数、列数和每个像素所占的位数,在进行 DICOM 图像显示时,需要把它们写入 DIB 信息头。

在显示医学图像时,还有一个问题值得关注。医学图像通常具有较高的灰度分辨率,比如 CT 图像、MRI 图像、数字化 X 胶片都达到 12 位,4096 灰度级。这些图像转化为 DIB 位图显示时,最大灰度级为 256 级,显然会造成图像的失真和细节的丢失,不能应用于医疗实践中。为了在普通显示器上显示医学图像的细节部分,可以采用加窗显示的方法。这里引入窗宽(Windows Width)和窗位(Windows Center)两个概念。窗宽是指要显示的图像数据的灰度范围,而窗位指这个范围的中心值。显示时把低于窗口下限的值置为 0,高于窗口上限的值置为 255,其他灰度级映射到 0~255 灰度级上,这样就可显示感兴趣的区域。DICOM 图像文件中(0028, 1050)、(0028, 1051)分别规定了窗位和窗宽,它们指示在灰度值的某一位置选取一定宽度来观察一些组织。

(2) DICOM 文件格式和其他格式文件的相互转换

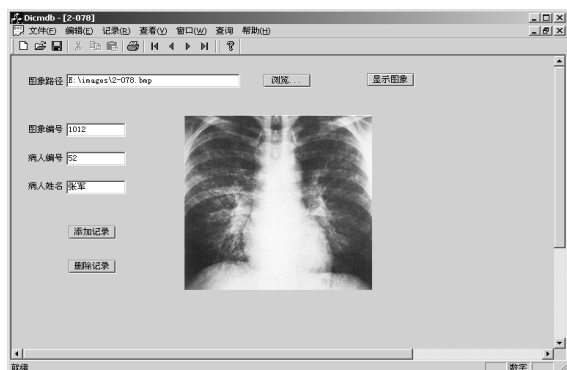


图 11.17 DICOM 图像的读取和转换界面

由于医院的医疗设备多种多样,产生的图像数据格式也不尽相同,为了使系统能够不受图像格式的限制,就要有专门的接口实现图像格式的转换。

在图像数据库系统内部一般采用 BMP、JPEG 等常见格式,这样做的目的是为了更方便处理和显示,所以系统要有将 DICOM 格式文件转化为 BMP 或 JPEG 等常见格式的能力;同时, DICOM 标准既是一种文件格式,又是一种通信标准,为了方便通信,系统又要能够将其他格式文件转化为 DICOM 格式。图 11.17 是一个典型的 DICOM 图像读取和转换界面。

11.4.6 小结

数字化医院是数字化城市极其重要的组成部分,而在城市建立数字化医院的进程中, PACS 的架构是一条必经之路。PACS 的建成将为电子病历和 HIS-RIS-PACS 集成系统的出现、远程医疗服务系统的实现打下基础。在这样投资和规模均较大的项目中,充分的技术论证、分步实施、务实求新、在社会效益和经济效益两方面得到可靠的保障,是确定方案时必须充分考虑到因素。

11.5 基于多分辨率分析的图像融合方法

11.5.1 图像融合的层次

一般情况下,图像融合由低到高分为 3 个层次:数据级(Pixel-level)融合、特征级(Feature-level)融合和决策级(Decision-level)融合。图 11.18 所示为像素级图像融合过程的示意图。像素级融合是在信息的最底层进行的,是将各幅源图像或者源图像的变换图像中的对应像素进行融合,从而获得一幅新的图像,其主要目的是图像增强、图像分割和图像分类,从而为人工判读图像或进一步的特征融合提供更佳的输入信息,目前大部分研究都集中在该层次上。像素级融合必须在严格的配准前提条件下,融合之前首先要对图像进行预处理的工作,包括降噪、几何校正、辐射校正等,如果图像具有不同的分辨率,在融合前还需要做相应的映射处理,像素级图像融合的融合准确度最高,是特征级和决策级融合的基础,是最重要、最基本的图像融合方法,有利于图像的进一步分析、处理与理解,进而提供最优的决策和识别性能,然而所需处理的信息量也最大,对设备要求较高。

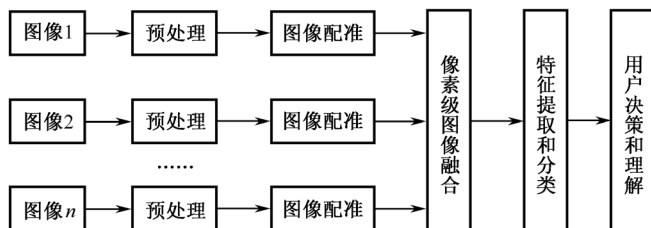


图 11.18 像素级图像融合过程

11.5.2 图像融合的方法

目前,国内外在像素级图像融合领域的研究方法大致可以分为两类:基于非多分辨率分析的图像融合方法和基于多分辨率分析的图像融合方法。

1. 基于非多分辨率分析的图像融合方法

基于非多分辨率分析的图像融合,一方面包含了传统的简单图像融合方法,如线性加权平均法、IHS (Intensity Hue Saturation) 变换融合法、PCA (Principal Component Analysis) 变换融合法及高通滤波法 (HPF) 等;另一方面也包含了智能图像融合方法,如基于神经网络、模糊理论的融合等。

简单的图像融合方法不对参加融合的源图像进行任何变换或者分解,而是直接对其提取的像素进行选择、平均或加权平均等简单处理后合成一幅融合图像。从实际应用效果来看,线性加权平均融合降低了图像的对比度;IHS 变换融合容易扭曲原始的光谱特性,产生退化现象;PCA 变换融合要求被替换与替换的数据之间有较强的相关性,而通常情况下,这种条件很难成立;HPF 融合在对高分辨率波段影像滤波的同时,也滤掉了大部分的纹理信息。因此,很多情况下,这些简单的图像融合法无法获得令人满意的效果。智能图像融合方法虽然是图像处理智能化发展的必然趋势,能够较好地解决图像融合过程中各个层次上出现的不完整性、不精确性、非结构化及建模困难等问题,但实现较为复杂。

2. 基于多分辨率分析的图像融合方法

基于多分辨率分析的图像融合方法是目前像素级图像融合技术中应用非常广泛且极为重要的一类算法,近年来,逐渐成为研究热点。由于其融合过程是在不同尺度、不同空间分辨率、不同分解层次上进行的,因此与以往的融合方法相比,基于多分辨率分析的图像融合方法可以获得效果明显改善的融合结果,同时具有适用场合广的特点。基于多分辨率分析的图像融合方法发展到现在,基本上先后经历了 3 个阶段,即基于金字塔变换、小波变换及超小波变换的图像融合法。其中,超小波分析是近来人们为改变小波分析的不足,常用基于小波技术基础之上的系列变换,即 Curvelet、Ridgelet、Contourlet、Bandelet、Beamlet 和 Wedgelet 等变换的统称,也有人称之为 X-let。

(1) 基于金字塔变换的图像融合方法

基于金字塔变换的图像融合是多分辨率图像融合的一个重要分支。其基本思想是:对每一幅源图像进行金字塔分解,然后通过从原始图像金字塔选择系数来构成融合金字塔,再将融合金字塔进行反变换即可得到融合图像。传统的高斯金字塔是一个在尺寸上逐层减半的一组图像序列,序列中的每一级图像均为其前一级图像低通滤波后作隔行隔列降采样,如图 11.19 所示。



图 11.19 Lena 图像的 4 层高斯金字塔变换

1983 年, Burt 基于图像压缩,在图像高斯金字塔的基础上,首次提出了拉普拉斯金字塔概念,并成功应用于图像融合。1989 年, Toet 利用类似的融合规则,在 HVS 对对比度敏感这一事实的启发下,利用比

率低通金字塔实现图像的融合,它能提供更好的细节信息给观察者。同年,Toet用数学形态学算子代替原框架中的高斯线性滤波,进行融合,所得到的图像金字塔序列被称为数学形态学金字塔。1993年,Burt和Kolczynski基于活性测度和匹配测度的融合规则,提出了梯度金字塔的融合算法。它是在图像高斯金字塔分解的基础上,对各层(除最高层外),用4个不同方向的梯度滤波算子,进而得到的梯度金字塔。

中值滤波是一种有效的非线性滤波工具,可将中值滤波和金字塔方法结合起来运用到图像融合中,对金字塔各层细节图像的融合,采用基于局部邻域平均梯度的融合规则,能有效地应用于多聚焦图像的融合,且具有强鲁棒性的优点。选用两幅多聚焦图像进行融合,中值滤波窗口和局部邻域窗口大小都选为 3×3 。该方法与基于拉普拉斯金字塔、对比度金字塔、梯度金字塔及小波变换的方法进行了对比,结果如图11.20所示,各种金字塔分解及小波分解层数均为3层。

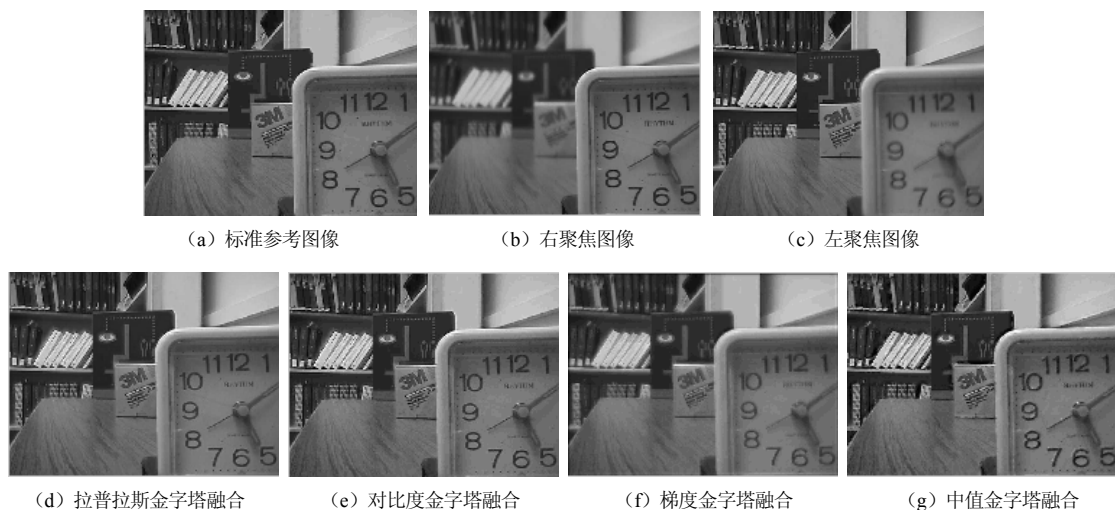


图 11.20 不同方法融合效果比较

(2) 基于小波变换的图像融合方法

1993年,Ranchin和Wald提出了第一个基于小波的融合系统,使用局部邻域窗口内的小波系数最大值作为活性测度,用绝对值取大法作为融合规则。1995年,Wilson等结合HVS的频率响应特性,把离散小波变换用于图像融合,并把小波分解后的小波变换系数的加权和作为活性测度。实验证明,用这种方法能获得比比率低通金字塔更好的视觉效果。1995年,Koren等把具有平移不变和无混淆特性的可操纵小波变换用于图像融合,以局部导向能量作为活性测度。只有对应于频率带中的活性测度值最大的系数才作为融合系数,用于重构图像。2000年,有学者提出了基于对比度的小波变换融合法,用小波分解各层的低频方向和高频方向的系数比作为该点的对比度,进而建立图像的对比度多分辨率序列用于融合。

在图11.21中,小波变换融合法使用的是db1小波,且采用高频系数绝对值取大法的融合规则。可见融合的整体视觉效果分别优于两个聚焦图像。

随着小波分析理论的发展和成熟,学者们在越来越多地研究一些新的融合方法和规则,并逐渐把小波包变换,多小波变换的不同算法用于图像融合,得到了良好的融合效果。

(3) 基于超小波变换的图像融合方法

1998年,美国斯坦福大学的Candes E J在小波理论的基础上,在其博士论文中给出了一种适合表示各向异性的多尺度几何分析方法——脊波(Ridgelet)变换。脊波变换是一种非自适应的高维函数表示方法,对含直线奇异的多变量函数能达到“最优”的逼近阶。Candes和Donoho于1999年提出曲波

(Curvelet) 变换理论, 也就是第一代 Curvelet, 该理论由 Ridgelet 理论衍生而来, 是由一种特殊的滤波过程和多尺度 Ridgelet 变换组合而成的。Curvelet 相对于小波, 具有更好的方向辨识能力, 更适合表示图像的几何特征。2000 年, 法国学者 Pennec 和 Mallat 提出了 Bandelet 变换。Bandelet 变换提供了一种新的基于边缘的图像表示方法, 能自适应地跟踪图像的几何正则方向。Minh Do 在 2002 年提出了 Contourlet 变换理论, 该变换继承了脊波和曲波的各向异性尺度关系, 其特征是用类似于线段的基结构来逼近源图像。2005 年, Arthur L.da 等提出了非下采样的 Contourlet 变换 (Nonsubsampled Contourlet Transform, NSCT), 通过非下采样的金字塔结构和非下采样的滤波器组构造而成, 是一种平移不变的、局域的、多方向的多分辨率表示方法, 是平移不变的 Contourlet 变换。



图 11.21 小波变换融合效果

图 11.22 所示为一组红外图像 (见图 11.22 (a)) 和可见光图像 (见图 11.22 (b)) 的融合结果。由于可见光和红外传感器的成像机理不同, 前者主要根据物体的光谱反射特性成像, 能很好地描述场景环境信息, 而后者主要根据物体的热辐射特性成像, 能很好的反映目标人物的存在特性。因此, 从图 11.22 (c)、图 11.22 (d) 中看出, 低频系数采用平均法会降低融合图像整体的对比度, 尤其是目标人物的对比度。而图 11.22 (e) 在这一点上要明显优于其余二种方法。从人物、栅栏、远处山脉和屋顶可以看出, 该方法在有效保留可见光场景环境信息的同时, 能结合红外目标的特性, 提取出丰富的边缘细节信息, 体现了 NSCT “沿” 边缘特性的捕捉能力。该方法对于多聚焦图像的融合同样十分有效。

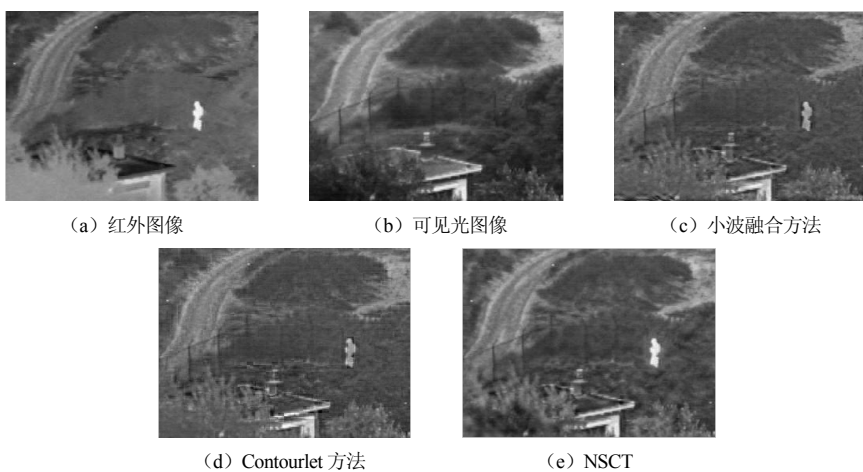


图 11.22 红外图像和可见光的融合结果

11.6 数字图像处理的发展趋势

数字图像处理技术是 20 世纪 60 年代初开始发展起来的, 经过了初创期、发展期、普及期及广泛应用几个阶段。经过几十年的研究与发展, 数字图像处理的理论和方法进一步完善, 应用范围更加广

阔,已经成为一门新兴的交叉学科,现已进入实用阶段。近几年来,随着计算机和各个相关领域研究的迅速发展,科学计算可视化、多媒体技术等研究和应用的兴起,数字图像处理从一个专门领域的学科,变成了一种新型的科学研究和人机界面的工具,其研究和应用呈现出蓬勃发展的崭新势头。数字图像处理的发展趋势主要反映在以下几个方面:

(1) 20 世纪末到 21 世纪初,在图像处理技术领域出现了一些新热点:在图像的安全技术方面,侧重研究图像数字水印和图像信息隐藏;在图像的自动识别方面,人体生物特征的提取与验证引起高度重视;在图像理解方面,图像匹配与融合等较高层的课题得到更多的关注;在图像数据库方面,对于基于内容的图像和视频检索的研究继续升温。

(2) “传统的”图像分割和图像压缩编码等领域仍有研究的价值。国内近十多年来占据文献最多的两个研究方向是图像分割和图像压缩编码。学者们对这些经典问题的研究兴趣反映了这个课题十分重要,如图像分割是由图像处理到图像分析的关键步骤,图像压缩编码具有很大的应用价值,同时也反映这些工作具有相当的难度和挑战性。

(3) 加强对人的视觉特性、心理学特性等的进一步研究。作为边缘学科,图像处理技术更应该注意借鉴其他学科的理论、技术和方法,完善图像处理的理论和技术体系。从人类视觉机理和机器处理技术两个方面去研究,将有效地推动图像处理技术的发展。

(4) 在硬件技术方面,要在进一步提高精度的同时着重解决处理速度问题。对于大数据量问题,可以采用并行多处理器技术。将软件实现用硬件实现,是解决速度问题的有效手段,即将图像处理的众多功能固化在芯片上使之更便于应用。值得一提的是,计算机技术也开始进入一个被称为后 PC 技术的时代。有学者认为,嵌入式智能工具将是 PC 和因特网之后最伟大的发明。现在嵌入式技术全面展开,已成为通信和消费类产品的共同发展方向。在通信领域,数字技术正在全面取代模拟技术。在广播电视领域,已开始应用于从模拟电视向数字电视的转变。市场前景十分看好的机顶盒的核心技术就是采用 32 位以上芯片级的嵌入式技术。嵌入式设备具有自然的人机交互界面,其以 GUI(图形用户界面)屏幕为中心的多媒体界面给人以很大的亲和力。小型手持嵌入式系统在手写文字输入、条码扫描等图像技术方面已取得初步成效。可以预见,嵌入式系统将在图像和视频处理方面发挥巨大的作用。

国家重点支持的高新技术领域包括以下图形和图像软件:支持多通道输入/输出的用户界面软件技术、基于内容的图形图像检索及管理软件技术、基于海量图像数据的服务软件技术、具有交互功能与可量测计算能力的 3D 软件技术、具有真实感的 3D 模型与 3D 景观生成软件技术、遥感图像处理与分析软件技术等。在医疗仪器技术、设备与医学专用软件方面支持医学影像技术:X 射线摄影成像技术(高频,中频)、新型高性能超声诊断技术(彩色 B 超)、功能影像和分子影像成像技术、新型图像识别和分析系统以及其他新型医学成像技术,包括电阻抗成像技术、光 CT 技术等。医学专用网络环境下的软件:医用标准化语言编译及电子病历(EMR)系统、电子健康档案系统、重大疾病专科临床信息系统、社区医疗健康信息系统以及实用三维数字医学影像后处理系统等。

(5) 图像处理技术在图像通信的研究和应用上具有广阔的市场前景。图像处理的发展将围绕 HDTV(高清晰度电视)的研制、图像及视频在网上的传输、因特网上的图像与视频检索、开展实时图像处理的理论及技术研究,向着高速、高分辨率、立体化、多媒体化、智能化和标准化方向发展。

(6) 智能图像信息处理新理论与新技术的研究。近年来,在图像处理领域引入了一些新的理论和技术,如小波分析、分形几何、人工神经网络、数学形态学、模糊逻辑、遗传算法、自适应信号处理技术、基于智能体的系统等。这些理论和技术使图像处理的手段更加智能化。

(7) 图像处理领域的标准化的研究。近 20 多年来,图像的存储、传输等方面已经建立了较多的国际标准,这些标准有力地推动了图像处理市场的发展。但图像信息的建库、检索和交流是一个重要的问题,有必要建立图像信息库,统一存放格式,建立标准子程序,统一检索方法。

11.7 基于 MATLAB GUI 图像处理软件的开发

MATLAB 提供了一个在图形用户界面 (GUI) 下制作图像处理软件平台的简便开发环境——GUIDE (Graphic User Interface Development Environment), 可以让开发者快速上手, 减少编程工作量, 提高开发效率, 有助于 MATLAB 程序的 GUI 集成。本节结合第 4 章例 4.6 实例简单介绍 GUIDE、GUIDE 的模板及其操作方法和图像处理软件平台的制作方法, 从而从感性上了解 GUIDE 的相关功能。对例 4.6 提出新的要求: ① 原始灰度图像的大小为 256×256 ; ② 符合条件 1 的图像可根据路径选择。

11.7.1 GUIDE 基本操作

(1) 在 MATLAB 命令窗口中输入命令行 `guide` 即可启动 GUIDE, 界面是如图 11.23 所示的 GUIDE Quick Start 对话框。

(2) 创建新的 GUI: 有 4 个 GUIDE 模板 (template), 默认选 “Blank GUI”, 或者打开已经存在的 GUI, 单击 OK 按钮, 打开版面设计工具。

(3) 将新建的 FIG 文件保存在希望的文件路径, 默认的 FIG 文件名为 `untitled.fig`。单击 OK 按钮, 进入下一个界面。

(4) 进入 GUIDE 编辑界面 (见图 11.24)。GUIDE 编辑界面主要由对象选择区、工具栏和布局区组成。其中左边的对象选择区包括 13 个对象。鼠标指向具体的对象, 系统会显示该对象名称的提示信息, 如 Push Button (按钮控件)、Axes (坐标轴控件)、Slider (滚动条控件)、Panel (面板控件) 等。

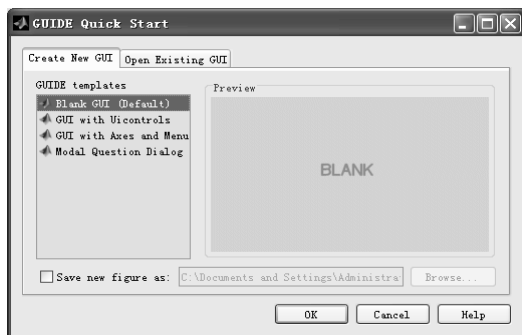


图 11.23 GUIDE 快速启动界面

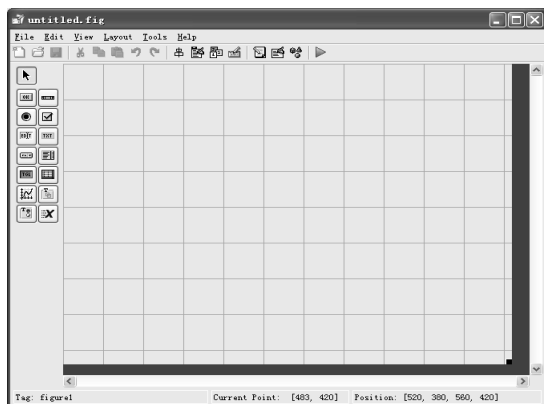


图 11.24 GUIDE 编辑界面

在菜单栏下面的工具栏包括 16 个按钮, 如 New Figure (新建 FIG 文件)、Run Figure (运行界面) 等。右下方的布局区用于安排 GUI 对象。使用 GUIDE 需要完成两项任务: GUI 对象属性设置及布局 and GUI 回调函数 (back call function) 编程。两个任务的结果分别存放在 FIG 文件和 M 文件中。

11.7.2 GUIDE 图像处理软件设计实例

利用 GUIDE 进行图像处理软件设计的步骤如下:

- (1) 对 GUI 对象进行布局。
- (2) 打开对象的属性查看器, 设置对象的相关属性。
- (3) 编写操作对象的回调函数。

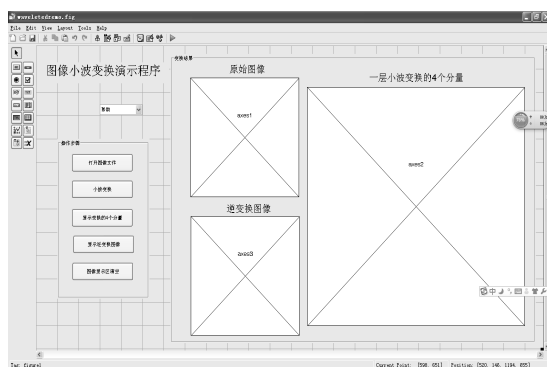
首先根据需求分析, 设计本软件所需要的对象及设想的布局 (见图 11.25)。



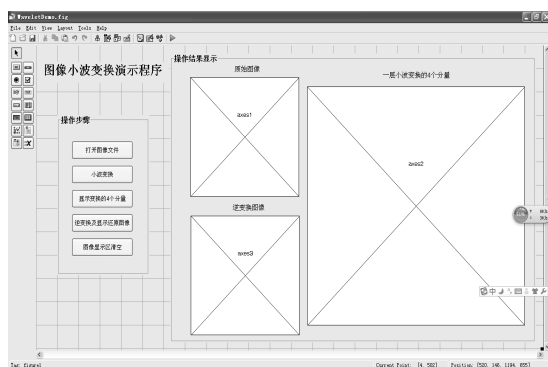
图 11.25 软件布局图

(1) GUI 对象布局。利用对象选择区的面板控件(用于“操作步骤”和“变换结果”)、按钮控件(用于“操作步骤”)、坐标轴控件(用于 3 个图像显示区)、静态文本控件(用于布局区必要的注释)、弹出式菜单控件(提供帮助功能),设计者直接用鼠标将所需的对象拖拽的目标位置即可完成布局设置,如图 11.26 (a) 所示。

(2) 设置对象的相关属性。为了保证界面美观和相应对象回调函数的编写,需单击右键打开各对象的属性查看器(Property Inspector)窗口,对相关参数进行设置。如第 2 个“pushbutton”按钮控件的 String 属性改为“小波变换”,Tag 属性改为“WaveletTransform”。控件的参数编辑完成后,如图 11.26 (b) 所示。运行并保存该 GUI,如图 11.25 所示的软件布局图。



(a) 设置控件后的效果图



(b) 控件属性修改后的效果

图 11.26 设置控件的效果图

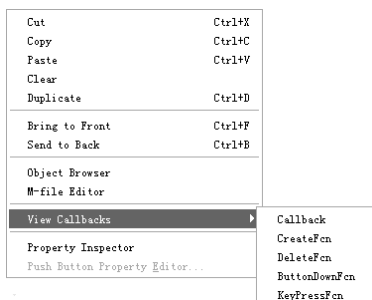


图 11.27 View Callbacks 菜单

(3) 编写操作对象的回调函数。如图 11.27 所示,单击右键,选择菜单命令 View Callbacks, 出现包含 5 种回调函数类型的子菜单。选择 Callback 命令,则打开 waveletdemo.m 文件并定位到控件所在的自动生成的回调函数体处,函数名为控件 Tag+Callback 类型名。GUIDE 仅给出了函数框架,开发者需要进一步补充实现回调函数相关功能的语句。

本软件需要编写的回调函数如下:

(1) GUI 初始化函数 waveletdemo_OpeningFcn

这里的功能主要是对坐标轴控件进行初始化操作。

```
% --- Executes just before WaveletDemo is made visible.
function waveletedremo_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject    handle to figure
% eventdata  reserved - to be defined in a future version of MATLAB
% handles     structure with handles and user data (see GUIDATA)
% varargin    command line arguments to WaveletDemo (see VARARGIN)

% Choose default command line output for WaveletDemo

handles.output = hObject;
I = ones(256,256); %供原始图像和逆变换图像的初始化
axes(handles.axes1); %坐标轴控件 1 处供显示原始图像
imshow(I); %显示初始化原始图像
WT = ones(512,512); %供小波变换四个分量的初始化
axes(handles.axes2); %坐标轴控件 2 处供显示小波变换四分量
imshow(WT); %显示初始化小波变换四分量
axes(handles.axes3); %坐标轴控件 3 处供显示逆变换图像
imshow(I); %显示逆变换图像

% Update handles structure
guidata(hObject, handles);
```

(2) “打开图像文件”按钮回调函数 `openfile_Callback`

通过以下的语句将 `mat` 文件转换成 `bmp` 格式文件: `load wbarb; I = ind2gray(X,map); imwrite(I, 'wbarb.bmp')`。

```
% --- Executes on button press in OpenFile.
function OpenFile_Callback(hObject, eventdata, handles)
% hObject    handle to OpenFile (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles     structure with handles and user data (see GUIDATA)
[file path] = uigetfile('*.bmp;*.jpg;*.png','请选择一幅图像');
if file == 0 warndlg('您得输入一幅图像');
%警告对话框提示输入合法图像文件
else I = imread(fullfile(path,file));
    axes(handles.axes1);
    imshow(I);
    handles.I = I;
end
% Update handles structure
guidata(hObject,handles);
```

(3) “小波变换”按钮回调函数 `wavelet_Callback`

```
function WaveletTransform_Callback(hObject, eventdata, handles)
X = double(handles.a);
    %对 X 进行二维小波变换, bior3.7 是双正交样条小波对应的滤波器
[cA1,cH1,cV1,cD1] = dwt2(X,'bior3.7');
A1 = upcoef2('a',cA1,'bior3.7',1);
H1 = upcoef2('h',cH1,'bior3.7',1);
V1 = upcoef2('v',cV1,'bior3.7',1);
D1 = upcoef2('d',cD1,'bior3.7',1);
    handles.A1 = A1;
handles.H1 = H1;
handles.V1 = V1;
```

```

handles.D1 = D1;
nbcol = 255;
%cod_X = wcodemat(X,nbcol);
cod_cA1 = wcodemat(cA1,nbcol);
cod_cH1 = wcodemat(cH1,nbcol);
cod_cV1 = wcodemat(cV1,nbcol);
cod_cD1 = wcodemat(cD1,nbcol);
dec2d = [cod_cA1, cod_cH1; cod_cV1, cod_cD1];
handles.dec2d = dec2d;
% Update handles structure
guidata(hObject, handles);
warndlg('小波变换完成!');

```

(4) “逆变换及显示还原图像”按钮回调函数

```

function DispInvImg_Callback(hObject, eventdata, handles)
% 进行小波变换的逆变换
A1 = handles.A1;
H1 = handles.H1;
V1 = handles.V1;
D1 = handles.D1;
Y = IDWT2(A1,H1,V1,D1, 'bior3.7');
%显示逆变换图像, 为了进行比较, 调整其尺寸和灰度范围与原始图像一致
Y1 = imresize(Y,0.461)
I = handles.I
Y2 = double(min(min(I)))+double((max(max(I))-min(min(I)))/...
(max(max(Y1))-min(min(Y1)))*(Y1-min(min(Y1))));
axes(handles.axes3)
imshow(Y2,[]);

```

(5) “逆变换及显示还原图像”按钮回调函数

```

function ClearDisp_Callback(hObject, eventdata, handles)
I = ones(256,256);
axes(handles.axes1);
imshow(I);
WT = ones(512,512);
axes(handles.axes2);
imshow(WT);
axes(handles.axes3);
imshow(I);

```

图 11.28 显示了 2 个测试图像进行小波变换的效果。



(a) 对 wbarb 图像的小波变换结果

(b) 对 testpad1 图像的小波变换结果

图 11.28 控件属性设置后的效果

本软件还有很大的改进空间,如适用不同的图像尺寸类型、选用不同的小波基、建立帮助功能、变换性能评估等。图 11.29 所示为 SBS 改性沥青专用图像处理与分析软件的 GUI 界面的截屏。

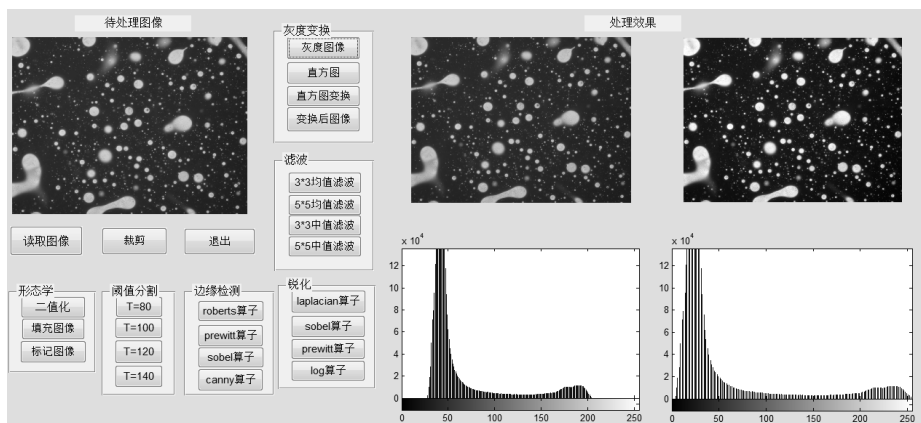


图 11.29 实际图像处理 GUI 界面

11.8 实验：基于 GUIDE 的图像处理软件开发

1. 实验目的

- (1) 学习使用 MATLAB GUIDE 的基本操作。
- (2) 以图像增强为例完成基本图像处理功能的回调函数编写。

2. 实验主要仪器设备

- (1) 台式机或笔记本电脑。
- (2) MATLAB 软件 (含 GUIDE 开发环境)。

3. 实验原理

实验步骤参考本章 11.7 节“基于 MATLAB GUI 图像处理软件的开发”,算法设计参考第 6 章“图像增强”,主要基于以下 4 种技术并将这些技术集成到一个 GUI 软件中。

- (1) 基于灰度直方图变换的图像增强。
- (2) 灰度修正图像增强。
- (3) 图像平滑滤波。
- (4) 图像锐化处理。

4. 实验内容

- (1) 直方图的计算及用直方图均衡原理增强图像。
- (2) 线性灰度变换图像增强。
- (3) 对受椒盐噪声污染的图像采用低通处理模板去噪。
- (4) 对受椒盐噪声污染的图像采用中值滤波去噪。
- (5) 对图像采用梯度算子和拉普拉斯算子进行锐化。

5. 实验步骤

- (1) 参考 11.7 节“基于 MATLAB GUI 图像处理软件的开发”,建立开发环境。

- (2) 参考第6章“图像增强”的算法编写相关回调函数。
- (3) 不断调试、优化，获得较满意的人机交互效果。

6. 实验报告要求

- (1) 说明实验原理。
- (2) 回答思考题。
- (3) 心得和体会。

7. 思考题

- (1) 结合教材第6章介绍的图像增强编程方法，说明采用 GUIDE 环境开发在此方面具有的特点。
- (2) 如果要对打开包括菜单栏和一些坐标轴图形对象进行 GUI 编程，应该选用什么 GUIDE 模板？
- (3) 如何在有限的屏幕空间合理设置控件对象？与显示器的分辨率有关吗？

本章小结

数字图像处理的应用十分广泛，本章主要从功能上介绍了数字图像处理的部分应用系统，包括图像处理在数字水印上的应用、基于数学形态学的图像颗粒度分析系统、数字化医院中的图像存档与通信系统、基于内容的图像检索系统、图像融合技术等。读者应该能够通过这些实例，举一反三，学习和提出其他的图像处理系统。

我们也可以从不同的应用领域讨论图像处理的应用，如在航空航天技术中的应用、生物医学中的应用、军事和公安、现代教育技术中的应用等。本章也介绍了数字图像处理的发展趋势，应该引起读者在学习和研究方面的关注。

本章花了一定的篇幅通过实例介绍基于 MATLAB GUI 图像处理软件的开发技术，旨在希望读者开动脑筋，积极探索，针对具体的图像处理系统，提出该系统的基本需求和技术解决方案，达到精于应用的学习的目的。

思考题与习题

11.1 从以下的专题进行讨论：图像处理在航空航天技术中的应用、图像处理在生物医学中的应用、图像处理在军事中的应用、图像处理在农业中的应用、图像处理在交通监控中的应用、图像处理在机器人技术中的应用、图像处理在虚拟现实中的应用。

11.2 从基本功能、系统结构等方面，构思在你身边需要的一种图像处理系统。

11.3 在 MATLAB GUIDE 环境下编制实现常用的图像分割功能，包括阈值分、区域生长等。

11.4 如果要对打开对话框编辑器进行 GUI 编程，应该选用什么 GUIDE 模板？

附录 A 常用词汇中英文对照表

数字

24-bit color 24 位颜色
2-D barcode 二维条形码
3-D image 三维图像

A

aberration 像差
Accelerated Graphic Port (AGP) 加速图形端口
accelerator board 加速卡
acoustic imaging 声学成像
acquisition 采集
adaptive thresholding 自适应阈值
additive color mixing 相加混色法
additive noise 加性噪声
affine transformation 仿射变换
air photo 航片
albedo 反照率
algebraic operations 代数运算
aliasing 混叠
amplitude 幅度
animation 动画
Analog to Digital Converter (ADC) 数模转换器
Application Programming Interface (API) 应用程序接口
Application Specific Integrated Circuit (ASIC) 专用集成电路
area function 面积函数
arithmetic coding 算术编码
array 数组
Artificial Intelligence (AI) 人工智能
Artificial Neural Network (ANN) 人工神经网络
autocorrelation matrix 自相关矩阵

B

B ultrasonic diagnostic instrument B 型超声诊断仪
back call function 回调函数
background 背景
bandpass filter 带通滤波器

bandstop filter 带阻滤波器
basis function 基函数
basis image 基图像
Bayes decision rule 贝叶斯判据
big data 大数据
binary image 二值图像
bit map 位图
bit per pixel (bpp) 位/像素 (像素深度单位)
blind image restoration 盲图像复原
blurring 模糊
bottom-hat transform 低帽变换
boundary tracing 边界跟踪
brightness 亮度
brightness modulation 辉度调制

C

Canny edge detector 坎尼边缘检测算子
Calibration 校准
Cathode Ray Tube (CRT) 阴极射线管
centralized spectrum 中心化频谱
chain code 链码
Charge Couple Device (CCD) 电荷耦合器件
Charge Injection Device (CID) 电荷注入器件
chromaticity diagram 色度图
classical restoration filters 经典复原滤波器
classification 分类
cloud computing 云计算
closing 闭 (运算)
CMYK space CMYK 空间
code position 码位
code-block 编码块
codebook 码本
color constancy 颜色恒常性
color histogram 颜色直方图
color image processing 彩色图像处理
Color Management System (CMS) 色彩管理系统
color map 颜色映射表

color model 颜色模型
 color space 彩色空间
 color table 颜色表
 color vector 颜色向量
 color wheel 颜色轮
 colorimetry 色度学
 command-line interpreter 命令行解释器
 Commission Internationale de L'Eclairage (法) 或
 International Commission on Illumination (英) (CIE)
 国际照明委员会
 Compact Disk (CD) 光盘
 complementary color 补色
 Compressed Sensing (CS) 压缩感知
 compression ratio 压缩比
 Computed Tomograph (CT) 计算机断层摄影
 Computer Graphics (CG) 计算机图形学
 computer vision 计算机视觉
 connectivity 连接性
 constrained restoration 约束复原
 Consultative Committee for International Telegraph and
 Telephone (CCITT) 国际电报电话咨询委员会
 Content-Based Image Retrieval (CBIR) 基于内容的图像
 检索
 Content-based Multimedia Information Retrieval (CMIR)
 基于内容的多媒体信息检索
 Continuous Wavelet Transform (CWT) 连续小波变换
 contour 轮廓
 contour coding 轮廓编码
 convolution 卷积
 convolution filtering 卷积滤波
 convolution operation 卷积操作
 correlation 相关
 criterion function 准则函数
 Curvelet Transformation 曲波变换

D

data-flow diagram 数据流图
 deconvolution 去卷积
 degradation 降质 (退化)
 dehazing 去雾
 denosing 去噪
 depth image 深度图像
 depth of field 景深
 diagonal detail 对角线细节
 Differential Pulse Code Modulation (DPCM) 差值脉冲
 编码调制
 digital camera 数码相机
 digital filter 数字滤波器
 digital hospital 数字化医院

digital image processing 数字图像处理
 Digital Imaging and Communication in Medicine
 (DICOM) 医学图像成像与通信标准
 Digital Light Processor (DLP) 数字光处理器
 Digital Micromirror Device (DMD) 数字微镜反射器
 Digital Negative (DNG) 数字负片格式
 digital projector 数码投影机
 Digital Scanning Converter (DSC) 数字扫表转换器
 Digital Single Lens Reflex (DSLR) Camera 数码单镜
 头反光照相机 (单反相)
 digital video camera (DV) 数码摄像机
 Digital Video Disk (DVD) 数字视盘
 digital watermark 数字水印
 digitizer 数字化器
 dilation 膨胀
 direct coding 直接编码
 Discrete Cosine Transform (DCT) 离散余弦变换
 Discrete Fourier Transform (DFT) 离散傅里叶变换
 discrete image transform 离散图像变换
 discrete K-L transform 离散卡胡南-列夫变换
 discrete mathematics 离散数学
 Discrete Wavelet Transform (DWT) 离散小波变换
 discretization 离散化
 distortion 失真
 distributed network 分布式网络
 dithering 抖动法
 domain block 域块
 Doppler effect 多普勒效应
 Dot Per Inch (DPI) 点/英寸
 drawing 图形
 duality 对偶性

E

edge detection 边缘检测
 edge linking 边缘连接
 eigenvalue 特征值
 eigenvector-based transform 基于特征向量的变换
 electronic image tube camera 电子成像管摄像机
 Electronic Shutter (ES) 电子快门
 Electronic Toll Collection (ETC) 电子不停车收费系统
 Enhanced Parallel Port (EPP) 增强型并行端口
 entropy 熵
 entropy coding 熵编码
 erosion 腐蚀
 estimation 估计
 exposure 曝光

F

false color 假彩色

Earth Observation System (EOS) 地球观测系统
 Fast Cosine Transform (FCT) 快速余弦变换
 feature 特征
 feature database 特征数据库
 feature matching 特征匹配
 fidelity 保真度
 fingerprint 指纹
 film scanning 胶片扫描
 FIR filter 有限冲激响应滤波器
 flash memory 闪存
 flood fill algorithm 泛洪填充算法
 Focus of Attention (FOA) 注意的焦点
 forth generation (4G) mobile communication 第四代移动通信
 fish eye lens 鱼眼镜头
 Fourier transform 傅里叶变换
 fractal coding 分形编码
 frame-store memory 帧存储
 frequency domain 频率域
 frequency response 频率响应
 full HD movie 全高清电影
 fuzzy set theory 模糊集理论

G

gamma correction 伽马校正
 gamut 色域
 Gaussian noise 高斯噪声
 geometric calibration 几何校正
 geometric distortion 几何失真
 geometric operation 几何操作
 global thresholding 全局阈值化
 gradient operator 梯度算子
 gradient-based segmentation 基于梯度的分割
 granularity function 颗粒度函数
 Graphic User Interface Developm Environment (GUIDE) (MATLAB) 图形用户接口开发环境
 Grassman's law 格拉斯曼定律
 grating 光栅
 gray image 灰度图像
 gray level 灰度
 gray-level correction 灰度修正
 gray-level histogram 灰度直方图
 gray-level interpolation 灰度插值
 gray-level transform 灰度变换
 gray-scale 灰阶

H

Haar transform 哈尔变换
 Hadamard transform 哈达玛变换

half toning 半调
 half-resolution 解析度减半
 High Definition Multimedia Interface (HDMI) 高清多媒体接口
 High Definition Television (HDTV) 高清晰度电视
 high-frequency enhancement filter 高频增强滤波器
 High Pass Filter (HPF) 高通滤波器
 High Resolution (HR) 高分辨率
 Hierarchical Data Format (HDF) HDF 格式
 histogram 直方图
 histogram equalization 直方图均衡化
 histogram specification 直方图规定化
 histogram transform 直方图变换
 hit 击中
 hole 孔洞
 hologram 全息图
 homomorphic enhancement 同态增强
 homomorphic filter 同态滤波器
 horizontal detail 水平细节
 Hotelling transform 霍特林变换
 Hough transform 霍夫变换
 HSV space HSV 空间
 hue 色调(色相)
 Huffman coding 哈夫曼编码

I

icon 图标
 image acquisition 图像采集
 image analysis 图像分析
 image classification 图像分类
 image coding 图像编码
 image communication 图像通信
 image compression 图像压缩
 image cropping 图像裁剪
 image database 图像数据库
 image description 图像描述
 image digitalization 图像数字化
 image digitizer 图像数字化器
 image enhancement 图像增强
 image fusion 图像融合
 image inpainting 图像修复
 image processing software 图像处理软件
 Image Processing Toolbox 图像处理工具箱
 image reconstruction 图像重建
 image representation 图像表示
 image restoration 图像复原(恢复)
 image retrieval 图像检索
 image segmentation 图像分割
 image structure 图像结构

image tiles 矩形图像片
 image transformation 图像变换
 image understanding 图像理解
 image-processing system 图像处理系统
 Image Processing Toolbox (IPT) (MATLAB) 图像处理工具箱
 imaging device 成像器件
 imaging system 成像系统
 index image 索引图像
 Information Appliance (IA) 信息电器
 integer-to-integer wavelet transform 整数小波变换
 intelligibility 可懂度
 intensity 亮度
 Internet of things (IOT) 物联网
 interpolation 插值
 interval 间隔 (区间)
 inverse color 补色
 Inverse Color Burn: 反选颜色加深; 反向色彩加深
 Inverse Color Dodge: 反选颜色减淡; 反向加亮色彩
 inverse filtering 反向滤波器法
 inverse transformation 逆变换
 invisuality 不可见性
 Iterated Function System (IFS) 迭代函数系统

J

jaggies 锯齿
 Jet Propulsion Laboratory (JPL) 喷气推进实验室
 Joint Bi-level Image Coding Experts Group (JBIG)
 JBIG 标准
 Joint Photographic Expert Group (JPEG) 联合影像专家组

K

Kahunen-Loeve Transform (KLT) 卡胡南-列夫变换
 Kirsch's operator 柯次算子

L

$L^*a^*b^*$ space $L^*a^*b^*$ 空间
 label matrix 标记矩阵
 Lagrange multiplier 拉格朗日乘子
 lasso 套索
 layer 图层 (Photoshop)
 light intensity 光强度
 Laplacian operator 拉普拉斯算子
 Least Signification Bit (LSB) 最低有效位
 LED panel LED 显示器
 lens 透镜
 Light Emitting Diode (LED) 发光二极管

light sensor 光传感器
 light valve 光阀
 linear algebra 线性代数
 linear algebraic restoration 线性代数复原
 linear system 线性系统
 linear transformations 线性变换
 Liquid Crystal display (LCD) LCD 显示器
 Local thresholding 局部阈值法
 Look-up Table (LUT) 查找表
 lossless compression 无损压缩
 lossy compression 有损压缩
 lossy image coding 有损图像编码
 Low Pass Filter (LPF) 低通滤波器
 Low Resolution (LR) 低分辨率
 Lux (lx) 勒克斯

M

machine vision 机器视觉
 maker 标识符
 Markov process 马尔可夫过程
 Markov Random Field (MRF) 马尔可夫随机场
 Marr's operator 马尔算子
 mask 掩膜
 mathematical morphology 数学形态学
 matrix 矩阵
 Maximum Entropy (PE) 最大熵
 maximum likelihood algorithm 最大似然算法
 Mean Opinion Score (MOS) 平均判分
 median filtering 中值滤波
 medical image 医学图像
 menu-driven interface 菜单驱动用户界面
 microscopic image 显微图像
 Minimum Mean Square Error (MMSE) 最小均方误差
 mirror 镜像
 miss 击不中
 model-based coding 基于模型的编码
 Moderate Resolution Imaging Spectroradiometer (MODIS)
 中分辨率成像光谱仪
 Modified DCT (MDCT) 改进的离散余弦变换
 Modified Huffman Coding (MHC) 改进的哈夫曼编码
 monochrome image 单色图像
 Monte Carlo restoration 蒙特卡罗复原法
 normalization 归一化
 motion 运动
 Moving Picture Expert Group (MPEG) 活动图像专家组
 multimedia 多媒体
 Multi-Resolution Analysis (MRA) 多分辨率分析

Multimedia Content Depiction Interface (MCDI) 多媒体内容描述界面

multispectral image processing 多光谱图像处理

Munsell model 孟塞尔模型

N

National Television System Committee (NTSC) 美国国家电视制式

Nearest Neighbor Rule (NNR) 最近邻准则

neural network (NN) 神经网络

Newton-Raphson iteration 牛顿-拉弗森迭代法

noise 噪声

noise contamination 噪声污染

noise model 噪声模型

noise modeling 噪声建模

non-linear system 非线性系统

numerical 数字

Nyquist sampling 奈奎斯特采样

O

object 目标(对象)

object measurement 目标测量

objective function 目标函数

object-oriented 面向对象

Open Source Computer Vision Library (OpenCV) 开源计算机视觉库

opening 开(运算)

opacity 不透明

Optical Character Recognition (OCR) 光学字符识别

optimal linear filter 最优线性滤波器

order statistic filter 排序统计滤波器

outline image 概要图像

overflow 溢出

P

paintings 画

panel 面板

Partial Differential Equation (PDE) 偏微分方程

pattern recognition 模式识别

perspective 透视

Pulse Coding Modulation (PCM) 脉冲编码调制

permanent display 永久显示设备

Phase Alternating Line (PAL) 行相位交错制式

photoconductor 光敏电阻

photoemissive device 光电发射器件

photographs 照片

picture 图片

Picture Archiving and Communication System (PACS) 图像存档和通信系统

piecewise linear transformation 分段线性变换

pixel (picture element) 像素

Plasma Display Panel (PDP) 等离子显示器

point operation 点操作(运算)

Point-Spread Function (PSF) 点扩散函数

power spectrum 功率谱

pyramid structure 金字塔结构

predictive coding 预测编码

prefix code 前缀码

Prewitt's operator 普雷瓦特算子

primary color 基色

Probability Density Function (PDF) 概率密度函数

probe (形态学) 探针

progressive transmission 渐进传输

projector 投影仪

pseudo color 伪彩色

Q

quantization 量化

quantizer 量化器

R

range block 值域块

raster image 光栅图像

reconstruction 重建

rectangular wave transform 方波变换

region growing 区域生长法

Region of Interest (ROI) 感兴趣区域

remote sensing image 遥感图像

resolution 分辨率

RGB space RGB 空间

Single Lens Reflex (SLR) camera 单反相机

Roberts gradient operator 罗伯特梯度算子

Robison's operator 罗比逊算子

robotics 机器人学

robustness 鲁棒性

round 舍入运算

run-length 行程

Run-Length Coding (RLC) 行程长度编码

S

sampled data 采样数据

sampling 采样

sampling aperture 采样孔

saturation 饱和度

scanner 扫描仪
 scanning mechanism 图像扫描机构
 screen shot/snap/capture 屏幕截图/抓图
 self-adaptive mesh coding 自适应网格编码
 self-similarity 自相似性
 set theory 集合论
 Shannon coding 香农编码
 shape 形状
 sharpen 锐化
 Short Time Fourier Transform (STFT) 短时傅里叶变换
 Signal-to-Noise Ratio (SNR) 信噪比
 similarity measure 相似性度量
 sinusoidal transform 正弦变换
 size measurement 尺寸测量
 skeleton 骨架
 skip blank coding 跳跃空白编码
 slant transform 斜变换
 Small Computer System Interface (SCSI) 小型计算机系统接口
 smoothing 平滑
 smoothness-constrained least-squares filtering 平滑度约束最小均方滤波
 Sobel's operator 索贝尔算子
 solid-state cameras 固态摄像机
 sparse representation 稀疏表示
 spatial domain 空间域 (空域)
 spatial transformation 空间变换
 split 分裂
 standard deviation 标准差
 statistic coding 统计编码
 stereo vision 立体视觉
 storage media 存储体
 structuring element (SE) 结构元素
 subtractive color mixing 相减混色法
 Super-Resolution (SR) 超分辨率

T

tag 标记
 template 模板
 textural feature 纹理特征
 texture 纹理
 Thin Film Transistor (TFT) 薄膜晶体管
 thinning 细化
 three primary colors 三基色
 threshold 阈值
 top-hat transform 高帽变换
 transform basis function 变换基函数

transform coding 变换编码
 transform domain filtering 变换域滤波
 tristimulus values 三刺激值
 true color 真彩色

U

ultrasonic imaging 超声成像
 unconstrained restoration 非约束复原
 Universal Serial Bus (USB) 通用串行总线
 update 更新

V

variable-length code 可变长码字
 variance-covariance matrix 协方差矩阵
 veracity 精确性
 Very Large Scale Integrate Circuit (VLSI) 超大规模集成电路
 Vector Quantization (VQ) 向量量化
 vertital detail 垂直细节
 video 视频
 video camera 电视摄像机
 Video CD (VCD) 视盘
 video surveillance system 视频监控系统
 vidicon 光电摄像管
 virtual reality 虚拟现实
 volatile display 暂时显示设备
 volume 数据量
 volume visualization 体视化
 Voronoi cell 胞腔

W

Walsh transform 沃尔什变换
 wavelet series expansion 小波级数展开
 Wavelet Transform (WT) 小波变换
 Wiener filtering 维纳滤波
 window 窗口
 Windowed Fourier Transform (WFT) 加窗傅里叶变换
 WLAN 无线局域网
 wraparound error 交叠误差

X

X-ray imaging X射线成像
 xG 第x代

Y

YIQ space YIQ空间
 YUV space YUV空间

附录 B 常用 MATLAB 图像处理工具箱函数

表 B.1 图像显示

函数名	功能
colorbar	显示颜色条
colormap	建立当前图像的颜色查找表
getimage	从坐标轴获得图像数据
image	将矩阵显示为图像，可返回一个图像的句柄给一个 image 对象
imshow	显示图像
montage	在矩形框中显示多幅图像
immovie	创建多帧索引图的电影动画
subimage	在一幅图中显示多幅图像
subplot	将多个子图画到一个图上
trueSize	调整图像显示尺寸
warp	将图像显示到纹理映射表面
zoom	缩放图像

表 B.2 文件 I/O

函数名	功能
imfinfo	返回图像文件信息
imread	从图像文件中读取图像
imwrite	将图像写入到图像文件中
load	将以 mat 为扩展名的图像文件调入到内存
save	将工作空间中的变量保存到以 mat 为扩展名的图像文件中

表 B.3 代数操作

函数名	功能
imabsdiff	两幅图像的绝对差值
imadd	两幅图像的加法
imcomplement	补足一幅图像
imdivide	两幅图像的除法
imlincomb	两幅图像的线性组合
immultiply	两幅图像的乘法
imsubtract	两幅图像的减法

表 B.4 几何操作

函数名	功能
imcrop	在指定的矩形裁剪图像
imresize	使用指定的插值方法调整图像大小
imrotate	使用指定的插值方法按逆时针方向将图像旋转任意指定的角度
interp2	二维数据插值

表 B.5 像素和统计处理

函 数 名	功 能
col2im	重排矩阵列为图像块
im2col	重排图像块为矩阵列
corr2	计算两幅图像（矩阵）的二维相关系数
imcontour	创建图像数据的轮廓图
imfeature	计算图像区域的特征尺寸
imhist	计算并显示图像数据的直方图
impixel	确定像素颜色
improfile	沿线段计算剖面图的像素值
mean2	计算矩阵元素的平均值
mean	求向量平均值
median	求向量中值
numel	求图像像素的总数
pixval	显示图像像素的信息
rand	产生均匀分布的随机序列
randn	产生正态分布的随机序列
reshape	新排列矩阵
size	计算图像的大小
sort	对给定的向量或矩阵排序
sum	对数组元素求和
std	计算给定向量标准方差
std2	计算矩阵元素的标准偏差

表 B.6 图像复原

函 数 名	功 能
deconvblind	盲去卷积复原图像
deconvreg	规则化滤波器复原图像
deconvwnr	维纳滤波器复原图像
edgetaper	使图像边缘振铃逐渐减弱

表 B.7 图像分析

函 数 名	功 能
edge	识别灰度图像中的边界
qtdecomp	执行四叉树分解
qtgetblk	获得四叉树分解块值
qtsetblk	设置四叉树分解块值

表 B.8 图像增强

函 数 名	功 能
imadjust	对比度调整
histeq	直方图均衡
imnoise	给图像增加噪声，包括高斯白噪声、黑白像素点噪声、乘积性噪声
medfilt2	二维中值滤波器
ordfilt2	顺序统计滤波器
wiener2	二维自适应除噪滤波器

表 B.9 图像线性滤波及二维线性滤波器设计

函 数 名	功 能
conv2	二维卷积
convmtx2	计算二维卷积矩阵
convn	多维卷积
fspecial	产生预定义滤波器
filter	一维线性数字滤波, 通常与产生预定义的滤波器函数 fspecial 配套使用
filter2	二维线性数字滤波, 通常与产生预定义的滤波器函数 fspecial 配套使用
freqspace	确定二维频率响应间隔
freqz2	计算二维频率响应
fsamp2	用频率抽样设计二维 FIR 滤波器
ftrans2	用频率转换设计二维 FIR 滤波器
fwindl	用一维窗口方法设计二维 FIR 滤波器
fwind2	用二维窗口方法设计二维 FIR 滤波器
imfilter	对任意类型数组或多维图像进行滤波

表 B.10 图像变换

函 数 名	功 能
fft	计算一维快速傅里叶变换
ifft	计算一维傅里叶逆变换
fft2	计算二维 FFT
ifft2	计算二维逆 FFT
fftn	计算多维 FFT
ifftn	计算多维逆 FFT
fftshift	直流分量移到频谱中心
dct	计算离散余弦变换
idct	计算离散逆余弦变换
dct2	计算二维离散余弦变换
idct2	计算二维逆离散余弦变换
dctmtx	计算 DCT 变换矩阵
dwt	一维离散小波变换
idwt	一维离散小波逆变换
wavedec	一维信号的小波分解
waverec	一维信号的小波重构
appcoef	从一维小波分解的分解结构中提取信号的低频系数, 常和 waverec 函数配套使用
detcoef	从一维小波分解的分解结构中提取信号的高频系数, 常和 waverec 函数配套使用
dwt2	二维离散小波变换
idwt2	二维离散小波逆变换
wavedec2	二维信号的小波分解
waverec2	二维信号的小波重构
appcoef2	从二维小波分解的分解结构中提取信号的低频系数, 常和 waverec2 函数配套使用
detcoef2	从二维小波分解的分解结构中提取信号的高频系数, 常和 waverec2 函数配套使用
wcodemat	对数据矩阵进行伪彩色编码
radon	计算 Radon 变换

表 B.11 图像邻域操作与块操作

函 数 名	功 能
bestblk	选择块处理的块大小
blkproc	对图像进行分块处理
col2im	重排矩阵列成图像块
colfilt	使用列方向函数进行邻域运算
im2col	重排图像块为矩阵列
nlfilter	进行一般邻域运算

表 B.12 二值图像操作

函 数 名	功 能
applylut	使用查找表进行邻域操作
bwarea	计算二值图像中的目标区域
bweuler	计算二值图像的欧拉数
bwfill	二值图像背景区域填充
bwlabel	标识二值图像中的连接成分
bwconncomp	查找二值图像的连通分量
bwmorph	二值图像形态运算
bwperim	确定二值图像中的目标边界
bwselect	选择二值图像中的目标
imdilate	对二值图像进行膨胀运算
imerode	对二值图像进行腐蚀运算
imopen	对二值图像进行开运算
imclose	对二值图像进行闭运算
makelut	构造（applylut）函数使用的查找表

表 B.13 基于区域的图像处理

函 数 名	功 能
roicolor	根据颜色选择要处理的区域
roifill	在任意区域内平滑插值
roifilt2	对要处理的区域滤波
roipoly	选择要处理的多边形区域

表 B.14 图像颜色图操作

函 数 名	功 能
brighten	颜色图变亮或变暗
cmpermute	重新排列颜色图中的颜色
cmunique	寻找唯一的颜色图及相应的图像
colormap	设置或获取颜色图
imapprox	由较少颜色的图像近似索引图像
rgbplot	绘制 RGB 颜色图

表 B.15 颜色空间转换

函 数 名	功 能
rgb2hsv, hsv2rgb	RGB 颜色空间值与 HSV 颜色空间相互转换
rgb2ycbcr, ycbcr2rgb	RGB 颜色空间和 YCbCr 颜色空间相互转换
rgb2ntsc, ntsc2rgb	RGB 颜色空间和 NTSC（YIQ）颜色空间相互转换

表 B.16 图像类型转换

函 数 名	功 能
dither	通过抖动增加图像颜色分辨率
gray2ind, ind2gray	灰度图像转换和为索引图像相互转换
rgb2gray	RGB 图像转换为灰度图像
rgb2ind, ind2rgb	转换 RGB 图像和索引图像相互转换为索引图像
im2bw	通过阈值化方法将图像转换为二值图像
mat2gray	将矩阵转换为灰度图像
grayslice	通过阈值化方法将从灰度图像转换为索引图像
isgray, isbw, isrgb, isind	判断图像是否为灰度级图像、黑白二值图像、真彩色 RGB 图像、索引图像

表 B.17 数据类型转换

函 数 名	功 能
double	将无符号 8 位整数转换为双精度浮点数
im2double	把图像数组转换成 double 精度类型
im2single	把图像数组转换成 single 精度类型
im2uint16	将图象数组转换成 unit16 类型
im2uint8	将图象数组转换成 unit8 类型
round	对浮点数转换为整数
uint16	将双精度浮点数转换为无符号 16 位整数
uint8	将双精度浮点数转换为无符号 8 位整数

参 考 文 献

- [1] K. R. Castleman. Digital Image Processing (Reprint Edition) . Pearson Education North Asia Limited & Tsinghua University Press, 2003
- [2] R. C. Gonzalez, R. E. Woods. Digital Image Processing (3rd Edition). Pearson Education, 2010
- [3] Milan Sonka, Vaclav Hlavac, Roger Boyle. Image Processing, Analysis and Machine Vision (Second Edition). New York: An International Thomson Publishing Company, 1998
- [4] 胡学龙. 数字图像处理 (第二版) . 北京: 电子工业出版社, 2011
- [5] 阮秋琦. 数字图像处理学 (第二版) . 北京: 电子工业出版社, 2007
- [6] 章毓晋. 图像处理和分析教程. 北京: 人民邮电出版社, 2009
- [7] 贾云得. 机器视觉. 北京: 科学出版社, 2000
- [8] 张德丰. 数字图像处理 (MATLAB 版) . 北京: 人民邮电出版社, 2009
- [9] 王大凯, 彭进业. 小波分析及其在信号处理中的应用. 北京: 电子工业出版社, 2006
- [10] 崔屹. 图像处理与分析——数学形态学方法及应用. 北京: 科学出版社, 2000
- [11] 包尚联, 周明全. 医学影像处理与分析. 北京: 电子工业出版社, 2003
- [12] 吴乐南, 胡学龙, 等. 多媒体及其相关技术的原理与应用. 南京: 东南大学出版社/台湾儒林出版公司, 1995
- [13] 严立中, 胡学龙, 李志军. 现代声像技术 (第二版) . 北京: 电子工业出版社, 2009
- [14] 胡学龙, 李志军, 宋鸣. “数字图像处理”多媒体双语教学的探索. 电气与电子教学学报, 2002.10, 24(5): 31-33
- [15] 胡学龙, 沈洁. 利用分水岭分割的多分辨率遥感图像融合算法. 中国图像图形学报, 2010, 15(5): 820-823
- [16] 胡学龙, 王志坚. 基于分数低阶矩的图像阈值检测和颗粒度分析算法. 中国图像图形学报, 2008, 13 (10) : 1821-1824
- [17] 黄凯奇, 王桥, 吴镇扬, 胡学龙. 基于人眼视觉特性的多尺度彩色图像增强算法. 中国图像图形学报, 2003, 8(11): 1242-1246
- [18] 胡学龙, 程茜, 杨莉. 基于形态滤波和梯度模糊 GHMT 的组合边缘检测. 电子测量与仪器学报, 2008, 22(3): 55-58
- [19] 胡学龙, 陆峰. 一种基于第二代小波变换的图像压缩方案. 电子测量与仪器学报. 2003, 17(4): 37-42
- [20] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, Image inpainting, SIGGRAPH 2000. 2000: 417-424
- [21] Xuelong HU, Min ZHANG, Nan JIANG, Xiang YIN. An Algorithm for Microscopic Image Granularity Calculation Based on Mathematical Morphology. Journal of Computational Information System (JCIS). 2010,6(12): 4177-4183
- [22] Huimin LU, Xuelong HU, Lifeng ZHANG, Shiyuan YANG, Seiichi SERIKAWA. Local Energy Based Image Fusion in Sharp Frequency Localized Contourlet Transform. Journal of Computational Information System (JCIS). 2010, 6(12): 3997-4005
- [23] Xuelong HU, Wei TIAN, Yongai ZHENG. A Watermarking Algorithm Based on Block Energy Analysis of Wavelet Transform's Coefficients and Lorenz Chaotic Attractor. International Journal of Digital Content Technology and its Applications (JDCTA) ,2010,4(9): 161-167